

Last high-noon for UNESCO

The British government seems determined to pull out of UNESCO. Here is a last-minute plea that it should change its mind. But UNESCO's problems will remain.

THE fine Corbusier building in Paris inhabited by UNESCO (United Nations Educational, Scientific and Cultural Organization) is not noticeably upset by the prospect that the United Kingdom may cease to be a member at the end of the year (three weeks from now). After all, the least successful of the UN's agencies has just returned from a conference at Sofia at which it seemed to make concessions to its critics, and gained the impression that some of them had been silenced. After the departure of the United States (responsible for 30 per cent of the 1984 budget), which seems to have left most things unchanged, how can the defection of the United Kingdom (with its mere 7 per cent) shake the edifice substantially? The simple answer is that the present time is a lull before a storm. UNESCO is going to have to change, not merely because of the complaints against it but from necessity. The strongest argument why the British government should not now pull out is that, if it stays, it has a good chance of being showered with the plaudits that will be showered on the reformers.

A still better reason for staying in is that there is much to do. UNESCO's errors are mostly errors of commission. It has tilted at impossible windmills, such as the objective of bringing about a new world information order (a device for making sure that governments are never offended by what their newspapers publish). Yet UNESCO has also done good works, it remains, for example, the only substantial source of support for the International Council of Scientific Unions. By any standards, a programme composed of such a mixture of the impossibly utopian and the practically utilitarian cannot but be endearing. The challenge, and the task that will remain even if Britain leaves, is to hammer this programme into shape.

For all its outward toughness about UNESCO, even the British government has not faced the central error in UNESCO's way of doing business, the generally accepted belief that the organization is meant to play a decisive part in the development of developing countries. Mr Timothy Raison, the British Foreign Office minister responsible for overseas aid (and for UNESCO), who has been sympathetically hard-headed on British membership over the past year, fell into the familiar trap in his speech during the House of Commons debate on UNESCO last month. Get those civil servants out of Paris, and into the Field, is the cry. The reality is that there are half a dozen other UN agencies whose terms of reference enjoin them explicitly to assist development, and whose budgets are more nearly commensurate with the task. A large part of UNESCO's difficulty over the past few years is that the majority of its members (developing countries) have insisted that it should become another aid agency, and that its paymasters (governments like the British) have weakly fallen in with the proposal.

By all accounts, the same error was repeated at last week's meeting of the British National Commission for UNESCO, the representative body of the great and the good which the government feels compelled to consult (but not compelled to listen to) at times like these. Most voices spoke for UNESCO's further transformation into a lowly version of, say, the Rockefeller Foundation that would remain hamstrung by its constitution, and by the need to deal equally with all of its constituents. The need for development, as this year's experience in most of Africa has all too plainly shown, is too urgent to be safely

neglected for much longer. But that, unfortunately, does not imply that UNESCO, already driven far off course, should be thrown ineffectually into the battle. The scale of its operations is simply too small to make much difference. But equally, it is not necessary that UNESCO should continue indefinitely to be the chief source of support for worthy organizations such as the International Council of Scientific Unions, which could (and should) keep itself alive by means of contributions from its members. UNESCO needs most of all to discover what it is for, and to pursue those goals wholeheartedly.

What might be done? It is not so long since UNESCO claimed to be an important influence in the teaching of science and technology in schools. It retains some influence in that field, spending modest sums on useful good works such as the design and manufacture of school science equipment, notably through its centre at the University of Delhi. But UNESCO has never had the nous to appreciate the generality of the difficulties of science teaching, the simple truth that rich countries such as the United States are as perplexed as the poorest among the UN's members. That is a field that a reformed UNESCO would plough. And why not extend the same principles both to other parts of the curriculum of the young?

In short, UNESCO's programmes should be devised to span problems within its terms of reference (education, science and culture) which are common to its members, not the exclusive concern of the poorest among them (or of any other class). Among present programmes, that to catalogue and, if possible, to preserve sites of archaeological or cultural importance to the international community is a good model for what might be done (if a poor model for administration). In the past, UNESCO has similarly helped to reach compromises between the rich and poor countries on questions such as copyright, where the trick has been to reconcile conflicting interests. There are many other issues of this kind that need attention, but which will not be attended to if too many UNESCO members quit.

That is the prize the British government should stay to help to achieve. There are few who would not accept its view that UNESCO as it stands is a badly administered can of worms which, by its reputation, does more harm than good to the cause of international cooperation. The obvious difficulty is that, if UNESCO should now collapse, it would not be reinvented within the lifetime of anybody now alive. Yet the tasks that cry out to be tackled are too important to wait that long, while the shocks to UNESCO of the past two years provide an opportunity that will not soon recur. □

Changing of the guard

The Royal Society has a new president, but it may also need a new policy.

SIR Andrew Huxley, who retired last weekend as president of the Royal Society, has predictably done a splendid job during the past five years. Undemonstratively temperament, almost the opposite of a power-seeker, his capacity (and liking) for intellectual work has enabled Britain's best known society to enhance its reputation for giving advice that cannot be ignored. During his spell in office, Sir Andrew has also emerged (again



predictably) as a courageous president; he has told off his opposite number in Moscow for the Soviet treatment of Sakharov and, more recently (see p.399), the organizers of next year's archaeological congress at Southampton for their indifference to the principles of freedom in science. That the administration of the society has been refreshingly improved during his tenure of office is an unexpected bonus. He is affectionately forgiven for his intervention on behalf of the British Museum (Natural History) — which should find itself a better name — in its dispute with *Nature* over the public presentation of evolution (see *Nature* 291, 373; 1981).

Huxley's successor, Sir George Porter, will almost certainly do as well. His record in research is similarly distinguished (with a Nobel Prize, while they last, a qualification for the job among other things), but he also has a foot in another camp, the popularization of science. Especially as director of the Royal Institution (Humphry Davy's home for Michael Faraday), Porter has done more than most working professionals to make sure that the world at large should have a better understanding of why science is important. It is no surprise that he should have said at the weekend that, for the next two or three years, the public understanding of science will be one of the Royal Society's preoccupations. There is a great deal of enlightenment to be effected. Everybody will hope that the new president can work a little wonder.

The obvious difficulty is that this is only part of the battle that needs to be fought. Logic would suggest that if the British electorate were better informed about the excitement of what happens in laboratories, it would sense the potential economic as well as cultural value of what is called research and instruct their elected representatives accordingly; and the consequence of that, the argument goes, is that British governments would no longer be able to deal with budgets for research as if they were any old kind of public spending. Unfortunately, in this as in other fields, logic is an inadequate guarantor of good sense, and is in any case slow to work its way through the electoral system. And although most British scientists will now say that their troubles stem from the cuts which there have been in their budgets, that is only half the truth. It matters at least as much that the system of research organizations spending what funds there are has been slow to accommodate itself to the changing environment. The simple consequence has been that potentially creative people have too often allowed themselves to be persuaded that they cannot hope to accomplish much.

In these depressing circumstances, the most urgent need of the British research enterprise is for leadership. It will help a great deal if researchers discover the Royal Society to be hard at work instructing people at large that science is a worthwhile enterprise, but there is more that needs doing. In particular, there is a need that some organization that stands for science, but which also understands the government's need to spend less, should more openly function as an informed critic of mismanagement of the kind that has marked the past few years. Who better than the Royal Society?

There are three stock explanations, one of which is that the Royal Society depends on public funds for most of its support, another that the society delivers its opinions confidentially, to those (in the government and elsewhere) who must act on them and, third, that a society composed of individuals elected for their academic distinction cannot be expected to have what might be called a corporate view. There is something in each of these arguments. It is a particular embarrassment that the Royal Society should both be dependent on the Advisory Board for the Research Councils for its own funds and an *ex officio* member of the same committee, no doubt because of the potential value of its candour. The view that the giving of advice through private channels inhibits public criticism is less cogent. And while it would probably be impossible to win an agreed view by the Royal Society's members on, say, the best way to finance the British Broadcasting Corporation or even the wisdom of replacing *Polaris* by Trident submarines, there is every chance that they would unite behind a measured and public criticism of

government policy towards British science in the past five (or even fifteen) years. What seems not to be understood is the degree to which morale in the laboratories would be stiffened by the emergence of an influential monitor of the present course of black events. It is possible that Sir George Porter has just such a plan hidden behind the slogan of public understanding. Now, at least, the Royal Society has a policy unit that could be the source of the analysis needed to sustain such a role. Why not take it up? There is much to be gained and, now, not much to be lost. □

Broadcasting in chaos

The British government has encouraged a sense of being mean that now extends to the BBC.

LIKE public policy on strategic arms, public policy on broadcasting is also an issue with which the technical community must grapple. The simplest reason is simply given; without the technology, there would be no issue, so that the innovators have a continuing responsibility. The more complicated but more interesting reason is that the innovators (or the inheritors) have a proper self-interest in the use made of the technology; will it, they may ask, further our interest that there may be more discovery and innovation? Those who may be sympathetic to this notion should pay some attention to what is happening to the British Broadcasting Corporation (BBC), certainly the best broadcasting organization in the world.

The BBC is constitutionally a public corporation, legally given (by charter) the right to function as if it were a corporation (or joint-stock company, as the Victorians would have said) but also saddled with the responsibility to broadcast material that gives neither personal nor political offence. But unlike other corporations, the BBC raises the funds it needs to keep on the air by means of an annual payment by those who own television and radio sets. This arrangement, conceived of in the 1920s, when television had not been invented, and when even the ownership of a radio receiver was a privilege, persists to the embarrassment of both the BBC (which needs the money) and the British government (which must periodically sanction an increase in the licence fee as, grudgingly, it did earlier this year).

So should not the market play a more prominent part in matching public expectations of broadcasting organizations to what the broadcasters provide? Especially because it has not escaped public attention that most broadcasting organizations are financed on a commercial basis, selling part of the time for which their audiences are prepared to listen to outsiders with other messages to impart, should not the BBC also sell time to advertisers? It would be unjust to the present British government to complain that it has insisted that this question should be answered. It has merely created the climate in which the question cannot be avoided.

The outcome is a committee under Professor Alan Peacock now brooding on the future financing of the BBC. Last week, the committee organized a meeting at which people whose views are well known repeated them for a gathering of others whose views are equally well known. (There were journalists in attendance, in the gallery.) The most radical (and the best) of the free-market solutions to the government's problem about the BBC (due to Mr Peter Jay, once British ambassador in Washington) is that people should pay for what they receive from among what is broadcast as they do for telephone calls. The trouble is that that is also politically unworkable. Most voters who now say that they would prefer that the BBC should take advertising than that the licence fee should be increased might turn nasty if they thought they would have to pay by the minute not the year. Yet the BBC must somehow be taken off the backs of successive British governments, all of which have been prepared to take credit for its quality while resenting the need to approve new licence fees. Why not turn the corporation into a public foundation which, unlike the Public Broadcasting Service in the United States, would be financed by an endowment, not an annual subvention? □

Apartheid in science

Crunch still ahead for archaeology congress

THE planners of next year's World Archaeological Congress at Southampton (England) have during the past week encountered mounting but conflicting pressures over the exclusion of South African scientists. The most ominous development is the decision of the International Union of Prehistoric and Protohistoric Sciences (IUPPS) to hold a special meeting of its executive committee (in Paris on 17 January 1986) to consider the matter. Meanwhile, the presidents of the Royal Society and the British Academy have spoken out against the exclusion, but the British committee organizing the congress has reaffirmed its earlier decision to ban South Africans.

The meeting of the executive committee of IUPPS (not to be confused with the British committee responsible for the congress) could affect what happens next year in Southampton. In principle, the committee could decide to withdraw its *imprimatur* from the congress. In that case, according to Professor John Evans, the president of IUPPS who is also chairman of the British committee, it would be necessary to "consider again" the arrangements for next year's congress.

The intervention of the presidents of the Royal Society and of the British Academy first took the form of a letter to *The Times*. On 17 November, Sir Andrew Huxley and Sir Randolph Quirk, speaking for their respective academies, referred to their "profound concern" at the decision to disinvite scientists from South Africa. They said that "it is an indispensable condition of holding an international conference that *bona-fide* scholars should be admitted irrespective of nationality, domicile or politics" and went on to describe the decision not to allow South Africans as "deplorable".

Huxley and Quirk went on to say that the decision could mean that Britain would cease to be regarded by bodies such as the International Council of Scientific Unions as "a fit place in which to hold an international scientific congress". Huxley expanded on this theme in his final anniversary address to the Royal Society last Saturday (30 November).

The meeting of the British National Committee on 20 November seems to have surprised those present by the weight of opinion in favour of continuing the ban. On one account, only one of a score of members spoke in favour of rescinding the ban. There are reports of overseas scientists having withdrawn from the congress, but the numbers are said to be "only about thirty". The possibility of moving the congress to another venue had been consi-

dered, only to be discounted on the grounds of lack of time.

At Southampton, the organizers have a sense of having being misunderstood, even misrepresented. Professor Peter Ucko, the secretary of the organizing committee, says that nobody has appreciated the trouble that had been taken to ensure the attendance of genuine scientists from developing countries. In this respect, he says, there is every prospect that the congress will be a success. Moreover, it had been arranged that participants would attend as individuals, not as nominated delegates of national organizations.

But, Ucko went on to say, people do not appreciate that "you cannot now hold a conference and invite people from South Africa if you also hope that there will be people from up to thirty Third World countries". This is the reality, Ucko says that conference organizers cannot ignore.

The sequence of events leading to the

ban on South Africans, beginning towards the end of the 1984-85 academic year, is not easily reconstructed. Ucko cannot at this stage recall whether the first representations came from the students at Southampton (whose union and, in particular, debating hall, is required for the congress) or from the local branch of the Association of University Teachers. The formal decision of the Southampton City Council to withdraw financial support came relatively late, he said, although there had been informal representations earlier in the summer.

Ucko, like many others involved in the conference organization, says that people have not paid enough attention to the moral indignation of those opposed to South African participation. The University of Southampton has not considered the trouble caused by its proposed hospitality for next year's conference, and may not do so. The vice-chancellor, Mr G.R. Higginson, said earlier this week that the university had no formal status in the congress except as the provider for accommodation during the university vacation. But Mr Higginson added that he had told the organizers during the summer that he saw no alternative to their decision to ban South Africans once the opposition to their presence had become apparent. □

Keyworth resigns as US science adviser

Washington

DR George A. Keyworth III, President Reagan's science adviser since 1981, announced last week that he is to resign from the position before the end of the year. Keyworth is widely credited with having successfully defended basic science even while other budgets were being cut, but also earned some unpopularity for

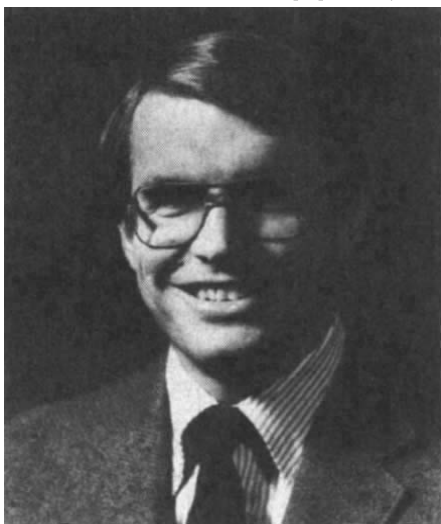
silence defences, which some fear could distort the pattern of financial support for basic research. But Keyworth said last week he was leaving at a time when everything was running smoothly: he was comfortable that SDI was on a firm footing and was encouraged by the general increased recognition of the importance of basic research, especially that conducted in universities.

William Carey, executive officer of the American Association for the Advancement of Science, said that Keyworth had earned a degree of access to the President that few other science advisers had achieved. He had also proven extremely effective in persuading the Office of Management and Budget to respect the importance of basic research, and had initiated important changes aimed at improving the efficiency of federal laboratories.

If Keyworth's term has been marked by generosity to basic science, his resignation comes at a time of widespread concern over the future level of support for science. Congress is considering legislation to reduce the federal budget deficit that would hit science hard, and many federal agencies have been told to expect budgets next year equal to or lower than this year's.

Keyworth is planning to establish a business with a former employee of the Central Intelligence Agency to advise companies on intelligence-gathering, where he is likely to command rather more than the \$70,000 salary he earned at the White House.

Tim Beardsley



being identified very much as a member of Keagan's team rather than as an emissary of the scientific community.

Keyworth has been a staunch supporter of President Reagan's controversial Strategic Defense Initiative (SDI) programme of research into anti-ballistic mis-

French research

Salomon speaks out (unpublished)

FRENCH technocratic centralism may have worked brilliantly for Ariane, nuclear power and high-speed trains, but it is not working, and will never work, for biotechnology, electronics and other consumer-oriented technologies. So argues Jean-Jacques Salomon, the only professor of science and technology policy in France, in a blistering — and unpublished — report now circulating at the Organisation for Economic Cooperation and Development (OECD), where Salomon was head of science for more than a decade.

Salomon's almost unrelieved criticism of French technology policy was commissioned 18 months ago by the prime minister, Laurent Fabius, in his then role of industry and research minister. Salomon submitted it to Fabius's successor, Hubert Curien, in July, but has heard nothing since. He is said to be furious that his passionate 160-page report may now simply gather dust.

The contents seem likely to guarantee that fate. For example, the report describes the creation of a grand ministry of research and technology by Jean-Pierre Chevènement as little more than "bureaucratic inflation". Scientists, writes Salomon, had much more influence over French policy in the pre-Chevènement days of the former and far smaller administration, the Délégation Générale de la Recherche Scientifique et Technique (DGRST).

The administrative consequences of Chevènement's expansion was a ministry split into several competing factions, says Salomon. This administration made only internal evaluations of its work, and rarely consulted its scientific advisers. Some advisory committees never met, according to the report, allowing bureaucrats to define programmes that were ends in themselves, not the means to innovation in the French economy.

In Salomon's report there are no scared cows, not even research spending. Basic research in France has enjoyed a 51 per cent budget rise since 1981 (without allowing for inflation), and applied programmes have seen a near doubling of spending. But according to Salomon, this attachment to a large increase in research and development spending as a fraction of gross national product has been "fetishism". (The ratio was 1.8 per cent in 1980, and 2.2 per cent now so a real increase has been achieved; the latest target is 3 per cent.) But, according to Salomon, the connection between research spending and economic growth is very loose: many other factors can matter more.

What the government should have been doing, according to the report, is actually to change the things that matter. But the author despairs whether in France anything can change. There have been 40 "re-

forms" of the French education system since 1945, Salomon counts, but none of them has finally changed anything. On the one hand, a true mass technical education remains out of reach despite frequent pious promises (of which the latest is Chevènement's, as education minister); and on the other, the top *grandes écoles*, the "engineering" schools that produce the elite of the French administration, in practice yield up graduates who "know nothing", neither science nor literature.

Also, Salomon suggests France must abandon its attachment to "great technological programmes" of state. These can only work when the state is, or controls, the purchaser; and can only fail when the purchaser is a multifarious public and the technologies various and rapidly changing (such as microelectronics). Here, the centralized French bureaucracy not only fails to be sensitive to the market but could not respond fast enough even if it were. The government should emphasize indirect action, where it has power and can change the climate, says Salomon.

• It is "indispensable" that there should be an independent, non-managerial, advisory unit to assess science and tech-

nology policy and ministerial action be established "as close as possible to the prime minister".

• The relevant ministries, with their competing and fragmented departments, should be streamlined, better directed to the marketplace.

• The "customer-contractor" principle, under which government departments should behave as customers for applied research by contract from research institutions, should be strictly applied to all big technology programmes.

• An external and public evaluation of the big technology programmes should be made "routine practice", as has become the case for the universities.

• The big programmes should take a lesser and lesser share of the cake, turning money towards "more supple and diversified" projects.

• Research and development aid to big industry must be analysed to determine if this is mere lame-duck support. Some 100 out of 1,500 French companies doing research receive 90 per cent of the government budget for industrial research and development, according to Salomon. There is danger of money going only to the strongest lobby, and more support is needed for innovative small companies.

• It is necessary to "rethink the French education system".

Robert Walgate

Cancer research

Ludwig branch lab to close

THE Ludwig Institute for Cancer Research is to close its London branch laboratories by the end of 1987. The unit, in collaboration with the Royal Marsden Hospital and the Institute of Cancer Research, is one of the major laboratories in the United Kingdom concerned with breast cancer. The Ludwig Institute has plans to open three more branches in London in the next two years and is negotiating with St Mary's Hospital, the Courtauld Institute and the Cardiothoracic Institute for sites. The staff at the present London laboratories in Sutton do not know whether they will be offered jobs at the new branches or even whether the organization intends to continue with breast cancer research. The reasons for the closure are obscure — clearly, lack of money cannot be one of them.

The Ludwig Institutes were set up in 1974 by Daniel Ludwig, one of the richest men in the world. There are at present nine branches, in the United Kingdom (2), Australia (2), Brazil, Belgium, Canada and Switzerland (2), and the organization plans to expand the number to 14. None of the branches are autonomous, but are administered by the Institute in Zurich. The present London laboratory was the first Ludwig branch to be established; its first director was Professor Munro Neville. Last year, Neville left to become a research administrator at the

organization's Zurich headquarters.

Staff at Sutton say that the post of director was not advertised after Neville left. The unit was without a director until the closure was announced, when Dr Tony Vickers, an ex-administrator from the British Medical Research Council, was appointed to oversee the rundown. Staff say that at the time of Neville's departure, the Ludwig executive assured them that the unit's future was not in doubt; eleven months later they received notice of closure. Although the technical staff have long-term contracts and are likely to be transferred, the scientific staff do not. And there is no indication that any of the new units will research breast cancer.

Scientific decisions by the Ludwig Institute are taken with the advice of a board of scientists from the United States and the United Kingdom. In the case of Sutton, staff are unconvinced that the normal procedure was followed and suspect that the decision was made at an executive level in Zurich. An explanation for the decision was promised for the staff by the end of November, but failed to materialize. Neville, speaking from Zurich this week, says that he "chose to leave" Sutton and that he knows nothing about the reasons for the decision to close it. And nobody else at the Zurich headquarters was prepared to explain the reasons for Sutton's demise.

Maxine Clarke

Embryo research

West German rules proposed

Bonn

THE long-awaited report of the West German interministerial commission on *in vitro* fertilization and related matters, set up in 1984 by the minister of research and technology, Heinz Riesenhuber, and the minister of justice, Hans A. Engelhard (FDU), was published last week (25 November). Part of the delicacy of the problem in West Germany arises because *in vitro* fertilization, even when carried out with the gametes of married couples, may be incompatible with the Basic Constitutional Law except when intended to lead to the birth of a child. The commission's proposals are to guide the drafting

of new legislation in the next few months.

The commission, chaired by the former president of the Constitutional Court, Ernst Benda, voted 17 to two for the majority opinion, which says that the use of donor sperm for *in vitro* fertilization may be permissible in special circumstances; the commission recommends that children conceived in this way should have a right to be told, at the age of 16, of the identity of the sperm donor.

In general, the report says, the technique should be used only for married couples who would otherwise be unable to conceive a child. The commission would strictly prohibit the use of *in vitro* fertilization for unmarried women, and the "rent a womb" practice in which fertile women bear a child for others for payment.

The most contentious of the recommendations are those concerned with the use of embryos in scientific experiments. In general terms, the commission would forbid embryogenesis for the purposes of scientific observation, but would allow state authorities the right to grant exceptions for investigations likely to benefit "the progress of medical science". The commission's findings have been criticized for not drawing a clear line. However the freezing of human embryos would be forbidden except in exceptional circumstances.

On the cloning of human embryos, the commission unanimously recommends a total ban, suggesting that such procedures should be made criminal offences. On gene therapy, the commission argues that attempts to alter the genetic constitution of somatic cells are in principle the same as organ transplants, and should therefore be allowed, but that gene transfer within the germ line should be prohibited.

One of the most controversial issues triggered by the appearance of the commission's report is that which would allow the analysis of the individual genome as when, for example, employers may seek to determine whether applicants for jobs are constitutionally unfit to do them. Although such checks are at present impracticable, the commission has been criticized by the trades unions and SPD for approving what may one day be possible.

One of the dissenting voices on the commission, that of Walter Doerfler, professor of genetics at the University of Cologne, argues that the commission's "basic attitude" towards new developments in human embryology is "very negative" and that emotions are a poor foundation for "hurried" legislation.

The other dissenting opinion is that of Professor Peter Petersen from Hannover, a psychologist. He is opposed to *in vitro* fertilization on the grounds that the physicians involved "do not really know what they are doing".

Jürgen Neffe

Danube dams

Austria takes responsibility

AUSTRIA is to take over responsibility for the construction of the controversial hydroelectric plant at Nagymaros, in Hungary, the Austrian finance minister, Dr Franz Vranitzky, announced last week. The cost, estimated at 8,000 million Austrian schillings (US\$450 million), will be borne by a consortium of Austrian commercial banks, and will be repaid, over 20 years, by electricity from the new power station. The deal, however, has been criticized by Austrian environmentalists, who feel that the contract unloads environmental problems on to Hungary.

The Nagymaros dam is the lowest stage in a joint Czechoslovak-Austrian peak-hour generating scheme. A storage reservoir at Dunakiliti, on the Danube below Bratislava, will, at peak hours, feed a generating station at Gabčíkovo, through a diversion channel that will take the main flow of the river away from the international border through Slovak territory. The Nagymaros dam, on the scenic Danube bend between Esztergom and Budapest, in addition to producing electricity for Hungary (and now, for Austria), will serve as a catchment for surges from Gabčíkovo.

Many Hungarian environmentalists have opposed the scheme, on the grounds that it will destroy the Danube wetlands which are the habitat of rare flora and fauna, ruin the amenity value of the bend (Budapest's chief resort area) and, by lowering the water-table, cause a pollution threat to the water supply of Budapest. The Hungarian government, the environmentalists urged, should either get Prague to agree to drop the scheme, or else withdraw unilaterally. Both courses, however, proved impractical politically, and in mid-August the Hungarian government reiterated its commitment to the scheme, though now with a postponed timetable and assurances of careful ecological monitoring.

Ironically, virtually the same environmental objections that the Hungarians had to the Dunakiliti-Gabčíkovo-Nagymaros installations were raised by the Czechoslovak government against Austrian plans to build a similar storage reservoir and power station at Hainburg.

In November 1984, the Czechoslovak foreign ministry announced that if the Hainburg dam were built, Czechoslovakia would demand massive compensation for ecological damage. A month later, after a long battle with the Konrad Lorenz Volksbegehren, Austria's main environmental organization, the Austrian government halted land-clearing operations at Hainburg, and, it now appears, began negotiating for a financial stake in Nagymaros.

Vera Rich

UK money for AIDS

MR Norman Fowler, Secretary of State for Social Services, this week announced that the British government would spend an extra £6.3 million on a package to combat the spread of acquired immune deficiency syndrome (AIDS). The money will be divided into five parts: £2.5 million for a national information campaign; £2.5 million for the Thames regional health authorities (75 per cent of the reported cases of AIDS in the United Kingdom are in London); £270,000 for six haemophilia counselling centres; £100,000 for training health professionals; and £750,000 for the Public Health Laboratory Service for testing blood samples for the AIDS virus.

The latest figures for the United Kingdom show that, of the 231 men and 10 women who have reported that they have the disease, 134 have died. A prediction of 1,837 new cases by 1988 was made at a meeting in London last week; the figure is higher than previous estimates because intravenous drug abusers are thought to be much more at risk than previously realized. The worry is that this group may be less amenable to public information campaigns than other "at risk" groups.

In a study published in the *Lancet* last month, it was reported that 38 per cent of serum samples from 106 Edinburgh drug abusers were positive for antibodies to the AIDS virus, with the proportions of men and women who registered seropositive about the same. Another report in the same issue analyses Centers of Disease Control data and shows that the relative risk of AIDS is four times greater for intravenous drug abusers than for homosexuals in the United States.

● The College of Health in London this week launched an addition to its Healthline service, which provides tapes of information on a range of medical subjects. Twelve new tapes give information on various aspects of AIDS; they can be heard by calling a London telephone number (1-980-4848) between 6 and 10 pm. Maxine Clarke

Polish universities

Government launches purge

At least 40 leading Polish university officials will lose their jobs as a result of the 1985 amendments to the Higher Education Act. The amendments increased the control of the Ministry of Science and Higher Education over the universities, and in particular, curtailed the rights of the universities freely to elect their own rectors, pro-rectors and deans. Despite assurances from government spokesmen last July that the changes then before the Sejm (Parliament) did not signal a purge in the universities, few academics were convinced. Their pessimism, it appears, was justified.

Under the amendments, rectors, pro-rectors and deans now in office will receive written confirmation of continued tenure from the ministry not later than the

end of January 1986, or else will be replaced. On 3 November, the underground newsletter of the pro-Solidarity Social Committee for Learning warned that the ministry planned to change the whole team of rector and four pro-rectors at Poznan's Adam Mickiewicz University.

This was surprising. Although the minister refused to confirm in office the rector elected in spring 1984, Professor Jerzy Fedorowicz, the current rector, Professor Franciszek Kaczmarek, a physicist, seemed an acceptable compromise.

Since then, however, there have been several pro-Solidarity incidents at the university. Moreover, the current Minister of Science and Higher Education, Dr Benon Miskiewicz, was the rector until his promotion to the post of minister shortly after the declaration of martial law. It seemed possible, therefore, that Minister Miskiewicz had singled out his own university as an example.

According to the Social Committee for Learning, the new rector was to be Professor Jacek Fisiak, a personal friend of the minister and a party hard-liner. Fisiak is a man of many parts. As head of the English department, he was responsible for the dismissal from the university, in 1978, of Dr Stanislaw Baranczak, the gifted poet and translator of English literature into Polish. He holds a British civil decoration, the Order of the British Empire, for his services to English studies in Poland.

Now, it appears that, according to Mr Andrzej Stolarski, chief spokesman for

the Ministry of Science and Higher Education, the new law specifically provides for a review of academic personnel throughout Poland, a statement somewhat at odds with the previous assurances. The Main Council for Higher Education, the supreme elected representative body of universities and higher colleges which, during the last academic year coordinated the campaign of the universities against the proposed revision of the law, met in closed session last Friday, but made no immediate statement. The Main Council, however, is known to be out of favour with the minister, who attacked it in parliament as being unrepresentative of the higher education community. It seems unlikely, therefore, that the Main Council will allow the dismissals to pass unquestioned.

Since the dismissals refer, so far, only to administrative posts, the rectors, pro-rectors and deans affected by the "review" will still, for the present, be able to carry on teaching — although there are strong rumours that the minister's next plan is for a "review" of teaching staff.

One rector who is to pay the price for putting his duty to his students before the demands of the authorities is Professor Wladyslaw Findeisen of Warsaw Polytechnic, who last year refused to identify and discipline those of his students who had carried the university banner at the funeral of the murdered Solidarity priest, Father Jerzy Popieluszko. The news of his dismissal last week evoked a massive and emotional demonstration by the polytechnic students, who gathered in the Great Hall of the Polytechnic. When the rector appeared, the students showered him with flowers.

Vera Rich

One US-Soviet pact

Washington

WHILE President Ronald Reagan and Mikhail Gorbachev were getting acquainted in Geneva the other week, another relatively unknown US-Soviet meeting was taking place in Moscow. Lee Thomas, administrator of the US Environmental Protection Agency (EPA), and Yuriy Izrael, chairman of the USSR State Committee for Hydrometeorology and Control of the Natural Environment, signed a memorandum presenting plans for cooperative work on 38 scientific and technical projects in the general field of monitoring the environment.

The meeting was the first high-level meeting on environmental cooperation between the two countries for six years. It took place under the auspices of a bilateral agreement originally signed in 1972 and since extended. Though the agreement has never been abandoned, deepening strains between the two countries had led to a steep decline in the number of scientists exchanged, from over 300 in 1977 to only 67 in 1984.

US officials say the United States has benefited in the past from the agreement, especially in earthquake prediction and environmental transport modelling. There have also been some difficulties: visas are often not available until the last minute, and last year one US official who had visited the Soviet Union several times under the agreement was suddenly denied a visa and slandered in *Izvestiya*.

At the recent Moscow meeting, eight areas of cooperation that had not been fruitful were abandoned, and four more were initiated. Nevertheless, the total number of scientific exchanges under the agreement is now expected to increase once more, according to Dr Gary Waxmonsky, EPA's executive secretary of US-Soviet programmes.

Tim Beardsley

Max Planck Gesellschaft

More cheer for Eureka than SDI

Bonn

THE Max Planck Society (MPG) seems eager that its 60 institutes and 10,000 employees should contribute to the European Eureka programme of technological research, but wary of the possibility that they might be caught up in the US Strategic Defense Initiative (SDI). So much emerged last week (27 November) when Professor Heinz A. Staab, the president of the society, introduced its annual report.

Professor Staab announced that the Max Planck Society is playing a leading part in one of the first Eureka projects, an inter-Europe collaboration in atmospheric chemistry. He said that the objective of the Eureka programme, at least as seen by MPG, is to complement the "existing one-way street to the United States" by more interconnections within Europe.

On SDI, Staab said that MPG has not so far been asked by the West German government to consider playing a part, so that MPG had not given formal consideration to the project. But he added that he could

not imagine that secret research could be carried out at any of the society's 60 institutes. He went on to say that plans for the development of basic research should not be "slipped over the scientific community" by politicians, whose influence should be used to strengthen existing collaborations.

MPG had funds of more than DM1,000 million to spend this year, and its budget of 1986 represents an increase of 3.8 per cent. But after allowing for the start-up costs of two new institutes, one for polymer research at Mainz and another for sociological research at Cologne, the increase for next year works out at only 2.8 per cent. Staab complained last week that the finance ministers of the *Länder* governments, which are responsible for half of MPG's budget, had originally worked for a mere 3 per cent increase of the total budget. Even now, Staab said, the increase would not be sufficient to cover the cost of inflation, reckoned to amount to between 5 and 6 per cent a year in the cost of research.

Jürgen Neffe

Planetary exploration

US Mars lobby gathers strength

Washington

THE "Mars underground", a network of enthusiasts dedicated to persuading the US government to support a manned mission to Mars, has again been raising its voice in public, pressing its case for a permanent base on Mars to the National Commission on Space. The signs are that it is being taken seriously.

The movement is made up both of university-based researchers and many in the National Aeronautics and Space Administration (NASA). As long ago as the early 1960s, NASA commissioned studies to investigate the feasibility of Mars missions, but enthusiasm waned as manned Moon landings became routine and the Apollo programme was brought to a premature close. But now the Mars underground seems to have found new strength.

Mr Leonard David, director of research of the National Commission on Space, is one of the biggest supporters of Mars ventures. He is to present a report on the future of the civilian space programme to the President and Congress by March 1986. The commission, appointed by President Reagan earlier this year, was explicitly told to take a bold long-term view, and it seems to be doing just that. It will look at options for the period to 2035 or even beyond, and will be taking a hard look at the Mars base plan as well as at proposals for a manned lunar base.

David points out that Mars has for many years been a vague long-term goal for NASA. With the shuttle now working at

least tolerably well and a space station apparently not too far off, the question that will arise is what they should be used for — whence the "case for Mars" group, which presented a 150-page summary of its plans to the commission last week at Stanford, California.

The group expects to use the first permanent manned bases on Mars established between 2010 and 2020; before then, though, bases may have been established on Mars's moons, Phobos and Deimos, which may contain enough water to use as a source of hydrogen and oxygen for fuel. The effort would be international and, after initial unmanned exploratory missions, the first crew of about 15 people would arrive for a ground spell of duty of up to two years.

The group lays much stress on using martian materials wherever possible, and notes that the martian atmosphere contains a high proportion of carbon dioxide, which could supply many essential needs for a permanent base.

The National Commission now has to weigh these ambitious plans against the rather less exciting lunar base plan. According to David, a lunar base might be a spin-off from a Mars base plan but is not generally thought to be an essential precursor. But there is one extra factor arguing for a Mars effort which could prove the most important of all: the Soviet Union has already spoken of plans for a manned Mars mission as early as the middle of next decade.

Tim Beardsley

Remote sensing

France puts trust in SPOT

SPOT 1, the first of two French observation satellites, is to be launched by Ariane on 11 January, only a few weeks late despite the failure of the last Ariane launch on 13 September. Arianespace, the commercializer of Ariane, will be crossing its fingers that adjustment to the design of a valve in the third stage will seal the leak of liquid hydrogen that caused the failure. So will the French space agency CNES, which has designed SPOT and which hopes it will steal the thunder from the US Landsat series of satellites.

CNES claims that twin high-resolution imaging instruments on SPOT 1, operating in the visible and near infrared, will be capable of identifying details down to about 10 metres, compared with the 30 metres of the last Landsat (Landsat 4) launched in July 1982).

CNES emphasizes, however, that the great innovation of SPOT is not so much its resolution as an ability to look away from the vertical. By this means, the satellite can in two passes some 90 minutes apart image the same Earth region from different angles, thus generating three-dimensional images. Simulations of such SPOT images have been made from aircraft, and they are certainly spectacular to view. These images should contain quite new cartographic information on vast regions of the more inaccessible parts of the Earth, and pose new challenges for information processing.

CNES has established a number of ground stations for receiving SPOT data, at which the data will be processed into directly usable images. In France, the principal station will be at Toulouse, where SPOT was designed and where all SPOT imagery is to be kept. A company called Spot Image has been established to sell and distribute images.

SPOT 1 will be accompanied into space, if all goes well, by the first Swedish scientific satellite, VIKING.

Subsequent Ariane launches are planned for February (two communications satellites, one for the US GTE Spacenet Corporation and one for the Brazilian government) and March (Intelsat V). Five additional launches are planned for 1986. Arianespace has also announced that despite its September setback, three institutions have since purchased future Ariane launches: one to orbit the first Luxembourg television satellite, which will broadcast to most of Europe, and one for a communications satellite for INMAR-SAT (the International Maritime Satellite Organisation).

These orders bring the total number of launch contracts of Arianespace to 37, of which 25 remain to be launched, the company says.

Robert Walgate

Parallel computers

GE aims at super-array contract

Schenectady, New York

GIANT General Electric (GE) is in the final stages of negotiating a contract with the US Department of Defense to complete the development and to begin the construction of what could be the world's most powerful computer.

The multimillion dollar project, a collaboration between GE's research and development centre here and the Artificial Intelligence Laboratory at the Massachusetts Institute of Technology (MIT), is for the construction of a machine with no fewer than 256,000 processors. Those responsible for the project wryly note that the new machine, called the Cross Omega Connection Machine, will have 100 times as much computing power as all the computers now installed at the centre.

The practical objective of the development is to construct a computer capable of processing in parallel the information in each picture element (pixel) of a standard 500 by 500 cell image. Suitably miniaturized, such a system could be carried by spacecraft or tanks and could make intelli-

gent decisions about the best response to incoming signals in video form.

The speed of the parallel array will be phenomenal — one million million operations a second. GE says that the architecture of the machine will be entirely novel, and that it will represent a milestone in the development of computing machinery. But the construction of the processors is also novel. By means of Very Large Integrated Circuit techniques, 64 processors will be incorporated on each of 4,000 semiconductor chips. By the use of optical lithography, GE claims that it now has in production at its plant at Research Triangle Park, North Carolina, integrated circuits with linewidths of only 1.2 micrometres, and that the research laboratory has made microcircuits with linewidths of 0.4 micrometres.

But Dr Phil Lewis, head of GE's computer science branch at Schenectady says: "We have to make a whole new computer science for this machine. And nobody knows how to do it."

Bill Johnstone

Japanese nuclear power

Straining at the US leash

Tokyo

JAPAN'S nuclear power industry, held back by US restrictions of fuel enrichment and reprocessing for more than a decade, has taken a great bound towards partial self-sufficiency in the production of its fuel, but probably in the wrong direction. Barely had the ink dried on an agreement with local officials to build a uranium enrichment plant based on centrifuge separation than the US Department of Energy sent shock waves through the Japanese industry by announcing that laser enrichment may be more economical. If this proves true, Japan's first commercial uranium enrichment plant due to open in 1991 will be outmoded and uneconomical.

Early this year, the governor of Aomori prefecture signed an agreement with Japan's Federation of Electric Power Companies for construction of a nuclear-fuel cycle complex in the prefecture on the northern tip of mainland Japan. It allows for plants for fuel enrichment and reprocessing as well as low-level waste storage. Although shaken by the US decision in June to switch to lasers, the Japanese government decided in October to press ahead with the enrichment plant next fiscal year, and a delegation was despatched to Washington in mid-November to seek revision of the 17-year-old Japan-US Atomic Power Agreement.

Behind these developments lies a 10-year struggle by Japan to break free of the US constraints which require "prior consent" before US fuel can be reprocessed. In 1972, Japan's Atomic Energy Commission recommended that commercial reprocessing plants should be developed after experience with a prototype plant, built by 1974 at Tokai-mura, north of Tokyo, by the Power Reactor and Nuclear Fuel Development Corporation (PNC), a semi-governmental research and development organization affiliated with the Science and Technology Agency.

With the tightening of US policy which the Indian nuclear explosion in May 1974, the start-up of operations at the Tokai plant was delayed by US restrictions until 1977, after Japan had signed the Treaty on Non-Proliferation of Nuclear Weapons. The delays by the Ford and Carter administrations raised bitter feelings, striking as they did at the heart of Japan's desires for energy self-sufficiency. PNC opened a centrifuge enrichment pilot plant at its Ningyo Toge works in 1979, and the steady trickle of fuel from this plant and the Tokai reprocessing plant has been used in PNC's prototype heavy-water reactor FUGEN and will be used in the prototype fast breeder reactor MONJU.

PNC has been less successful with its plans for radioactive waste storage. Last month it caused an uproar at Horonobe, a small town in the northern island of Hol-

laido, where it had planned a site for high-level storage with the approval of the town but in the face of opposition from the governor or Hokkaido, other local government leaders and residents. The Council of All-Hokkaido Labor Unions had mobilized a daily vigil by 300 members, reduced to 40 at weekends. PNC officials sneaked onto the site two weeks ago to choose sites for boreholes and for the location of a seismograph as well as to collect rock samples. A spokesman for the corporation said the inspection was carried out without notice to "avoid possible confusion among local residents".

Waste storage facilities will also be an integral part of the new complex to be built in Aomori. Around 66 million litres of low-level radioactive waste are at present stored at nuclear plants around the country. Earlier tentative plans to dump the waste in the Pacific from other Pacific countries, so that the Aomori complex will have storage space for up to one million 200-litre drums of waste in the first

stage and three million in the second have met with strong opposition. Proximity of the storage site on the Shimokita peninsula to the Misawa air base has raised fears of air-to-ground weapons going astray or crashing jets, but others favour the development as a stimulant to local economy.

Unlike the reprocessing and enrichment plants run by PNC, which are largely experimental in nature, the complex at Aomori will be a fully commercial enterprise backed by MITI and run by Japan Nuclear Fuel Industries Inc., a consortium established by the electric utilities and machinery companies in March this year. The plan is for Japan to produce domestically about 30 per cent of its enriched uranium needs by 2000. Japan would then be free to reprocess such "home-grown" fuel without prior consent from the US.

This panacea for Japan's dependence on energy imports may never materialize, however, if the laser-beam method of enrichment is successfully developed by the United States. Realizing this, the Japanese government announced in October plans to establish a joint government-private industry organization to develop the method.

David Swinbanks

Soviet *refusniks*

French scholarship in absentia

THE Paris-based Comité des Professions Médicales en faveur des Juifs d'URSS has established a bursary for *refusnik* scientists, taking the form of an invitation to study for a year in France. The award is to be called the Victor Hugo bursary to mark the appointment in 1882 of Victor Hugo as president of an emergency committee to help the Jews of the then Russian empire.

The first "recipient" will be Karen Khatchuryan, a 24-year-old ex-student of physics who was expelled from Moscow University in 1981 for "behaviour inconsistent with the name of a Soviet student" — in other words, for having filed an application to emigrate to Israel.

Whether the young man will be able to benefit from the award seems doubtful. If he was unable to obtain permission to emigrate to Israel, it seems unlikely that the Soviet authorities will let him go to France.

The bursary was announced during a meeting at the Paris Palais de Congrès devoted, in the first place, to medical problems relating to Soviet Jews. Some telling anecdotal evidence was presented on the decline in the proportion of Jewish school-leavers entering medical schools, particularly the distinguished establishments of Leningrad, Moscow, Kiev, Khar'kov and Vil'nyus.

Another panel dealt with the personal problems of the children of *refusniks*, forced to lead a double life between the Soviet culture they experience in school and the Jewish culture of their parents. Dr Albert Mallet, a psychiatrist who has vi-

sited the Soviet Union on several occasions, says that out of some 80 children examined, about half have developed serious personality disorders directly associated with their enforced lifestyle.

The third main panel at the meeting dealt with the political misuse of psychiatry. The case of Dr Zakhar Zunshain, a Jewish physicist from Latvia, suggests a change of Soviet practice. Undated copies of correspondence between Mrs Zunshain and the Deputy Minister of Health of the Latvian SSR, made public at the meeting, show that Dr Zunshain, a professor of physics who was sent to labour camp in March 1984 for having taken part in a demonstration outside the Bolshoi theatre in Moscow, had on some prior occasion been visited by a psychiatrist, at the request of the local police.

When Mrs Zunshain tried to obtain an explanation of this occurrence, the Deputy Minister, K. Berman, replied that it "should be considered quite normal". When denied a visa to emigrate, Dr Zunshain had refused to take "no" for an answer, and had continued to write "non-pertinent declarations which contained insults to the officials responsible", causing "serious doubts as to his mental health". The health law of the Latvian SSR, wrote Berman, provides for "prophylactic examination of citizens whose conduct and actions are characterized by a lack of critical spirit and opportunity", a ruling, which, if it were generally applied, could put the entire *refusnik* community at risk.

Vera Rich

Soviet seismology

Tajikistan earthquake reveals flaw

LAST month's earthquake in the Soviet republic of Tajikistan caused unnecessary damage, according to the local correspondent of *Izvestiya*, V. Surkov. The planning authorities of the Tajik SSR, he says, have for a long time "resisted" the classification of Leninabad *oblast'*, in the north of the republic, as liable to suffer earthquakes of force 8 on the Soviet 12-point scale. Not only had buildings been built to specifications suitable only for less seismic zones, he claims, but Leninabad city itself does not have a single seismic station.

Tajikistan is the most seismically active region of all the Soviet Union's 2 million km² of earthquake-prone terrain. The industrialization and development of the region has therefore called for considerable circumspection on the part of the geologists, and seismic risk maps for centres of population have been in existence for a number of years.



The latest forecast for Leninabad *oblast'*, which dates from 1982, increased the predicted maximum shocks from force 7 to force 8. The republic authorities, however, Surkov said, were reluctant to act on the new figures, and neither strengthened existing buildings nor modified the specifications for new ones. As a result, in the earthquake of 14 October, the towns of Kairakkum and Gafurov suffered heavy material damage.

Surkov's dispatch, from Leninabad *oblast'*, contradicts the impression given by earlier reports from the TASS correspondent, Aleksandr Nemirovskii, who filed from the Tajik capital of Dushanbe. According to Nemirovskii, "residential quake-proof houses built in recent years withstood the force-8 shock". This may be technically true — Surkov implies that some 10 per cent of housing was at least reparable — and Nemirovskii did go on to say that some people who had lived in old houses had had to be given shelter in tents, but the extent of the devastation, and the allegations of official negligence, come only from Surkov.

It is possible that even if the buildings had all been up to recommended standards, they might not have survived, since the shock in fact exceeded force 8. This means, Surkov notes, that future buildings in the area ought to be constructed to resist at least force 9. But some Soviet civil engineers seem doubtful about high-rise

construction in seismic areas, and after the Gazli earthquake of March 1984 it was decided to replace devastated apartment blocks with one-storey dwellings.

More significant, perhaps, is Surkov's claim that Leninabad *oblast'*, which accounts for one third of the population of Tajikistan, has insufficient seismic monitoring stations. Earthquake prediction is a major concern of Soviet geology. Recommended methods range from sampling the helium content of geothermal waters to observing the behaviour of local fauna.

The most reliable early-warning sys-

tem, however, if populations are to be evacuated and emergency services alerted, is the continuous monitoring of microtremors, which are then processed by computer to establish forecasts of the expected epicentre and strength of imminent earthquakes.

In Leninabad *oblast'*, there were not only too few monitoring stations but the communications links to the Central Asian Seismic Forecasting Centre in Dushanbe did not operate. The earthquake itself, moreover, damaged telephone lines to the devastated areas, so that emergency rescue services, while fully aware that their help was urgently needed, had no idea where to send it.

Vera Rich

Japan in chaos

Radical attacks halt railways

Tokyo

As in other advanced nations, the functioning of Japan's giant cities is very much dependent on fast urban transport systems and the sophisticated electronic communications that control them. Just how vital these communications systems are was shown when left-wing radicals, protesting at plans to privatize the national railway network, last week cut trackside cables in Tokyo and Osaka. Eleven million commuters were stranded and for half a day many businesses ceased to function.

Almost exactly a year ago, a small fire knocked out computer links and brought Japan's banking system to a halt (see *Nature* 312, 393; 1984). At that time, there was an outcry over the vulnerability of the emerging back-up systems. But it seems that ultra-left groups learned more from that incident than did the government.

Early on Friday morning, radical groups, armed with garden shears and knowledge of where to find the cables for the national railways' central traffic control and automatic train stop systems, struck at 23 locations in the Tokyo area. The first the railway authorities knew was when screens at the central computer headquarters went blank; when they tried to contact key stations through their own telephone network, that too was found to be dead. Similar events took place in Osaka. That morning, the two cities' entire public rail systems were paralysed.

Repairing the cables is a difficult task: a minimum of four hours is required when a single break is made in the more sophisticated communication links. When there are several cuts in the same cable, the difficulties are multiplied many times over. During the day, huge traffic jams developed and enormous queues formed at the stations of the private railway line still running. One queue in Tokyo was said to consist of 35,000 people. Businesses opened late and a fifth of Tokyo's schools closed for the day.

In a related incident, other radicals,

dressed in business suits and ties but wearing helmets, attacked a medium-sized station in the centre of downtown Tokyo with Molotov cocktails. They then flung away their helmets and disappeared among the crowds of commuters. The station burned throughout the day.

No statement has been issued by the attackers, but it is assumed that they were members of the People's Revolutionary Army of the ultra-left Chukaku-ha (Middle Core Faction). Their attacks were timed to coincide with a strike by a break-away chapter of the National Railmen's Union based in Chiba, a little to the east of Tokyo, protesting against plans to privatize the National Railways.

The Chukaku-ha is a strong and well-organized revolutionary group with probably around 3,000 members, many of them recruited at universities. Membership of its secret revolutionary army probably numbers several hundreds.

Previously, the group has been most conspicuous in its bitter fight against the construction of Tokyo's new international airport at Narita. Although the airport was opened seven years ago, the radicals are still fighting for it to be closed and for land expropriated from small farmers to be returned. The airport is constantly guarded by some thousand police, but this has not prevented frequent disruptions in service.

That radicals now seem willing to use similar tactics against the railways is bad news for the Japanese government. Railway authorities have already admitted that there is no way they can guard hundreds of kilometres of communication cables against attack, nor can they install a complete back-up system without ruinous expenditure. Selling off the railways is going to be a considerably tougher task with radical opposition which could continue indefinitely. And police fear that the Chukaku-ha may be planning something big for the Tokyo summit of world leaders next May.

Alun Anderson

British civil plutonium

SIR—Barnham, Hart, Nelson and Stevens (*Nature* 317, 213; 1985) conclude that at least 2.3 ± 0.8 tonnes of plutonium produced by the magnox nuclear stations of the UK generating boards cannot be accounted for, and may have been diverted to weapons use. An earlier version of this paper was submitted by the Campaign for Nuclear Disarmament (CND) to the Sizewell B Public Inquiry where it was debated at length. The Central Electricity Generating Board (CEGB) demonstrated that the analysis was in substantial error and no credence could be placed on its conclusions or, of course, the implications that the authors were striving to secure therefrom. While the revised paper takes account of some of the comments made by CEGB, the main criticisms still apply. These are as follows.

(1) The paper derives the estimate of 2.3 tonnes by attempting to calculate the total plutonium from the CEGB and South of Scotland Electricity Board (SSEB) stations since 1963, and compares this estimate of 47 tonnes with the amount the government has accounted for in parliamentary statements, augmented by the amounts the authors estimate to be in civil use in the United States, and to have been lost in reprocessing. Since the approach depends on estimating the difference between two large numbers, the conclusion is critically dependent on the accuracy of the calculations — clearly the calculation of plutonium production needs to be accurate to better than 5 per cent for the claimed discrepancy to be of any significance. It follows that a very careful calculation is necessary to obtain results of the accuracy needed, but although the paper claims an accuracy of ± 2 per cent, this is not substantiated. In particular, an accurate estimate is needed of the plutonium production as a function of irradiation; this requires calculations for each station using detailed data on station design and operation, and a theoretical model more complex than the simplified method adopted in the paper, which uses approximate data culled from a variety of sources. Since the three different methods of calculation described in the paper all use the same estimate of plutonium production as a function of fuel irradiation (the function $G(B)$), the three methods are not independent. The degree of agreement between their results provides no evidence on the accuracy with which this plutonium production function has been calculated.

(2) The authors attempt to validate their overall method of calculation by comparing estimates of plutonium production over a 6-year period with data provided by CEGB for amounts of plutonium in irradiated fuel despatched for reprocessing from each of our eight magnox stations. Table 4 of the paper concludes that this shows that their model underestimates

plutonium production by 3.7 per cent, implying that their estimate of 2.3 ± 0.8 tonnes for plutonium unaccounted for is an underestimate by $0.037 \times 47 = 1.7$ tonnes. However, as CEGB explained at the Sizewell Inquiry, this comparison is not valid because it does not allow for changes in the level of irradiated fuel stocks at the stations over the six-year period. The comparison is also selective since the paper reports the comparison for seven of the stations but excludes the data provided for the other station at Wylfa. If Wylfa is included, the results are substantially changed — the overall ratio of plutonium dispatches to calculated production changes from 1.037 to 0.948. If the authors had not excluded the data for Wylfa, they would presumably have concluded that the comparison shows their method overestimates plutonium production by 5.2 per cent and that the claimed discrepancy of 2.3 ± 0.8 tonnes is overestimated by 2.4 tonnes.

The board is not able to publish a definitive analysis of plutonium production for reasons that were explained at the Sizewell Inquiry. These relate to the period before 1971 when some of the plutonium produced, although itself remaining in civil use, was bartered with the United States for a quantity of highly enriched uranium which was used for defence purposes. It has been the policy of governments over the years not to disclose this figure, and we are also precluded from providing detailed information which would help others to derive it.

To sum up, the accuracy of the authors' calculations is not sufficient to support the allegation that plutonium produced in civil reactors in the United Kingdom may have been diverted to weapons use.

Repeated statements in Parliament by ministers have confirmed that no plutonium produced in CEGB reactors has been used for military purposes in the United Kingdom or exported for use in weapons. The government has also stated that it has no plans to use plutonium produced in the board's reactors for nuclear weapons and this is also the policy of the US government.

T. BROOM

(Director-General)

*Technology, Planning
and Research Division,
Central Electricity Generating Board,
Courtenay House, 18 Warwick Lane,
London EC4P 4EB, UK*

Free will

SIR—Nothing is gained by the invocation by D. H. Evans (*Nature* 307, 762; 1985) of the second law of hydrodynamics.

It may be true that free will implies the ability to think divergently; but determinism does not imply convergent thought: it implies only determinism. Given that the

thought processes of the scientist are the result only of causes in his environment, they can appear as convergent or divergent. They are convergent when the observer (who is also determined) perceives nothing new (in Evans' terms, no loss of entropy) which could not have been deduced logically from the observed facts. They are divergent when the observer perceives something new (either not deducible or apparently unrelated) which can subsequently be demonstrated by experiment. However, an observer with access to all the facts (that is, the whole internal and external environment of the scientist) might perceive the deductive logic of "divergent" thought.

Free will implies that the scientist has control over his divergent thought (which is intuitively absurd) while determinism implies only cause and effect. Convergence and divergence are then functions of the observer's environment, not the scientist's environment.

The entropy of the situation is more straightforward. Entropy is different in spatial and temporal regions of the Universe, and the variability might arise as a result of local fluctuations and aggregations. Given a spectrum of entropy, its value in one region is a statistical function (deterministically predictable but otherwise chosen (!)) and in any given region can be very high or very low. Thus in the spatial regions of living forms it can be very low: temporally it may continue to fall before it rises again.

Consciousness may indeed be a manifestation of biochemical processes, and the brain may function far from thermodynamic equilibrium: but that does not prevent it being determined in the same way that a whirlpool is determined.

S.J. STARKIE

*8 Weston Grove,
Fulbourn, Cambridge, UK*

Academic tenure

SIR—The question whether those who work as professors or researchers at academic institutions should enjoy security of tenure has been much discussed (see, for example, *Nature* 317, 464; 1985) but surely, in democracies, if job security is to be eroded for some, the same rules should apply to others. Yet in the armed services, those who graduate from military academies and similar institutions have tenure for the rest of their lives. They may be asked to retire early but are compensated with a pension, and at no point are they required to justify their promotion, largely automatic, by enumerating the battles they have won or the missiles they have fired accurately, the military equivalents of papers published or graduate students trained. Can this be fair?

J.M. GOLDSCHVARTZ

*Clavecimbellaan 273,
2287 VK Rijswijk zh,
The Netherlands*

Competition for big time opens

The launching of the Hubble Space Telescope next year will provide up to 3,000 hours of the best-ever telescope time. Here is how to bid for some of it.

THOSE hoping for a share of the time that will be available to users of the Hubble Space Telescope will at least now know to which address to write*. Towards the end of last month, the Space Science Institute on the Johns Hopkins campus at Baltimore put in the mail some thousands of copies of its formal *Call for proposals*, a document that explains the procedure applicants for telescope time must follow, together with more than a kilogram of supporting data, not to mention application blanks (twelve copies of which must be returned).

The documents give as full an account as there could be of how the potentially contentious issue of time allocation will be dealt with by the Space Science Institute. They are also, in passing, a vivid reminder of the potential power of the Hubble telescope, a 2.4-metre reflector equipped with instruments sensitive to radiation from the infrared (11,000 ångströms) to the ultra-violet (1,150 ångströms).

That the instrument will do more for observational astronomy than the building of the Hale telescope on Palomar Mountain in the 1930s seems generally accepted; it could yet outdo Galileo's contribution to the craft of observation. So competition for telescope time will be fierce.

So why will there be only 3,000 hours, rather less than 10 hours a day, of observing time each year? The new documents make plain the uncertainties that persist, the chief of which is that no date has yet been fixed for launching the telescope by means of the US space shuttle. And in the nature of things, nobody can guarantee that the instrument will find its way into orbit (at an altitude of 500 kilometres with an inclination of 28.5 degrees).

But the documents also explain that the efficiency with which the telescope can be reoriented towards novel targets is still an unknown quantity, whereas observations will not be feasible while the axis lies too close to the direction of the Sun, or the bright limbs of the outline of the Earth.

Applicants will also have to take account of the annoying complication of the South Atlantic Anomaly, the departure from regularity of the Earth's magnetic field that allows charged particles trapped in the magnetosphere to reach down

into the orbit of the telescope; the telescope detectors will be unusable not merely during passage through the anomaly but for an unquantifiable length of time thereafter, while the phosphorescence of the detectors makes the noise unacceptably high. One consequence is that the longest period of time when the continuous viewing of a target will be possible is estimated to be between 10 and 12 hours, a little more than at a high-latitude terrestrial observatory on a clear winter night.

The opportunities, on the other hand, are immense. The focal plane of the mirror is equipped with a variety of instruments, to each of which is assigned the radiation intersecting a predetermined patch of the plane, allowing the remainder of the radiation collected by the telescope to be gathered into the synoptic images of the wide-field camera (with the comparatively small field of view 154 arc-seconds square). This instrument will be able to acquire images, always in digital form, with a visual magnitude of 28. Perhaps more striking is the expectation that the use for astrometry of one of the three fine guidance sensors with which the telescope is equipped (two are needed for pointing the telescope) will allow angular positions to be determined to within 0.0016 arc seconds. This expectation, if borne out by experience, implies that it will be possible directly to determine the distances to important milestone stars used in constructing the distance scale of the Universe.

So how are these great opportunities to be shared? In principle, according to the Space Science Institute, anybody may apply for time. All proposals will be put through the same peer-review system, and time allocated on merit.

On the teasing question of deciding between the merits of proposals yielding quick spectacular results and long-term programmes bearing fruit only after the accumulation of a mass of data, the institute plans an arbitrary but workable solution. Each proposal will be considered by a disciplinary review panel, but final recommendations will be made by a Telescope Allocation Committee with general instructions to allocate roughly equal amounts of time to proposals in the categories small, medium and large. These allocations will account for roughly 90 per cent of the time, with the remainder reserved for allocation by the director, among other things so that unexpected opportunities may be seized and so that

the telescope may be used in novel ways.

The institute has already settled on three key projects to which special attention will be given, of which the chief is the determination of the distance scale, ideally so as to determine Hubble's constant to within ten per cent. The second consists of an invitation to exploit the potential of the wide-field camera so as to carry out a deep survey of selected regions of the sky, yielding information on such miscellaneous issues as the comet cloud beyond Neptune and the supernova rate in distant galaxies. Quasar absorption lines come third.

Speed will obviously be crucial. The documents make plain that applications for time will not succeed if the objective is a project identical with one already in the programme. Thus the first convincing proposal to measure Hubble's constant is likely to hold sway for two or three years. The other side of this coin of exclusivity is the institute's policy on the publication of data. At the outset, an observer will have exclusive access to the data gathered on his behalf for twelve months after the last of them have been stored, but thereafter, the data will become part of the general archive at Baltimore.

So how should intending applicants proceed? First, read the instructions carefully. The chances of being caught out by misunderstanding the technical performance of the instruments, or by miscalculating the length of the exposure required, are high. Those who succeed in the first round of applications will find themselves quickly asked for accurate positions of the guide stars they specify for their targets. Perhaps the best hope of winning time early is to go carefully through the list of projects put forward by those guaranteed exclusive use of the first six months (people who have helped to develop the telescope), looking for opportunities to make parallel observations of objects within the field of view of the Hubble telescope during that initial period.

But can it be worth the trouble when there is always a risk that the telescope may not be launched successfully? The risk of failure is no greater than that the shuttle itself might fail. But, whatever happens, the list of proposals on their way to Baltimore in the next three months is likely to serve future historians well as a record of the ambitions of the world's astronomers at this interesting time.

John Maddox

*General Observer Support Branch, Space Telescope Science Institute, PO Box 33139, Baltimore, Maryland 21218, USA.

Evolutionary biology

Sex ratios in wasps and aphids

from Robert M. May and Jon Seger

DARWIN seems to have framed most of the major problems in evolutionary ecology, and to have showed, at least in outline, how they might be solved. But he failed to crack the problem of the sex ratio. "In no case . . . would an inherited tendency to produce the sexes in equal numbers or to produce one sex in excess, be a direct advantage or disadvantage to certain individuals more than to others; . . . I formerly thought that when a tendency to produce the two sexes in equal numbers was advantageous to the species, it would follow from natural selection, but I now see that the whole problem is so intricate that it is safer to leave its solution for the future."¹ The future began in 1930 with R.A. Fisher; that it continues is shown by papers in *Science*² and this issue of *Nature*³.

Fisher pointed out that if we count grandchildren (rather than children), then it follows at once that there is an advan-

tage to producing the minority sex, whose members enjoy greater reproductive success, on average, than do members of the majority sex⁴. Thus males and females will be produced in equal numbers at the evolutionary equilibrium. If the cost of producing a male offspring differs from that of producing a female, then we need to state Fisher's result in its more general form, which says that at the equilibrium there is equal total investment in the two sexes. Darwin might easily have discovered this principle some 60 years before Fisher, had he thought to look ahead two generations instead of one when evaluating an individual's reproductive success.

Fisher's result assumes that there is population-wide competition for mates. In 1967, Hamilton⁵ showed that the ratio to investment may be dramatically female-biased under certain kinds of mating systems in which, owing to the physical

isolation of sibships, brothers compete to inseminate their sisters. Hamilton marshalled evidence demonstrating large female biases in a number of insects (mainly parasitoid wasps) that are likely to experience intense 'local mate competition' (LMC).

Hamilton's paper stimulated interest in many different kinds of problems concerning sex allocation. LMC itself has developed into a substantial industry, in part because its mathematics are seductively elegant, and in part because it explains patterns that do not have plausible alternative explanations based simply on the mechanics of sex determination. For example, it has often been claimed that the 1:1 sex ratios of mammals and birds are mere side-effects of the X-Y (or Z-W) chromosomal mechanism of sex determination. A recent book by Bull⁶ stands this reasoning on its head, by considering the evolution of a wide range of sex-determination mechanisms, taking into account their effects on the sex ratio.

Wasps have no sex chromosomes, being haplodiploid, so they can play

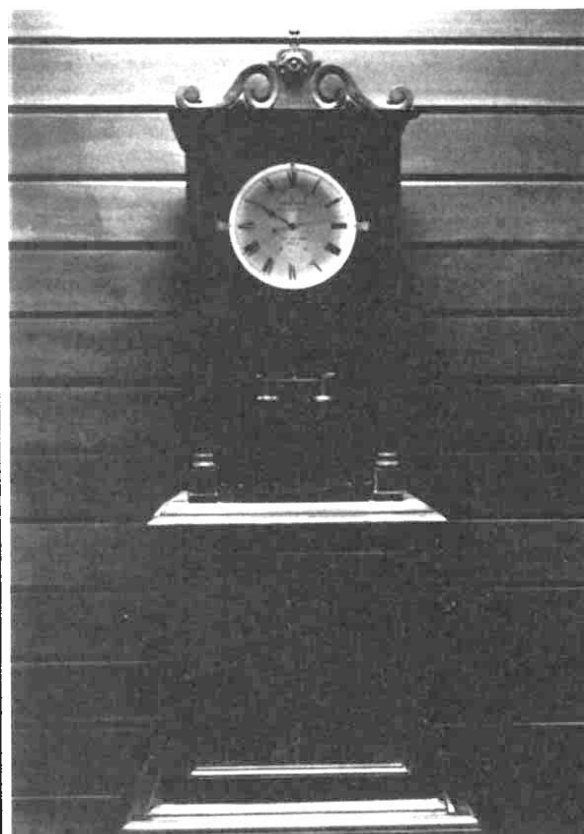
sophisticated sex-ratio games' involving extreme biases and, more interestingly, conditional responses to local variation in the environment or in the structure of the population. Theorists have explored the strategic possibilities at some length, and experimentalists have tested many of the resulting predictions using laboratory populations of the wasp *Nasonia vitripennis*, which has become the *Drosophila* of sex-ratio research. In no other field of evolutionary biology has it been possible to verify so many detailed quantitative predictions. Recent books by Charnov⁷ and by Trivers⁸ review the work on LMC and on many topics in the theory and practice of sex allocation.

Some of the most exciting recent advances have come not from the laboratory, but from the field. For several years there has been uncertainty as to whether the female-biased sex ratios seen under LMC are responses only to the competition for matings that occurs within groups of brothers, or whether they are also responses to the inbreeding with which such competition is almost always confounded. By comparing three species of fig wasps, Herre has recently shown that the biases reflect both the level of sib-competition and the average level of inbreeding⁹. Theoretical predictions, some derived for the first time by Herre, call for the more highly inbred species to show stronger female biases for any given number of foundresses in an individual fig (that is, for a given intensity of LMC). This is exactly what Herre sees in the three species. Not only are their sex ratios ranked in the expected order, but the actual data fall remarkably close to the theoretical curves relating sex ratio to foundress number. Herre's test of the subtle distinction between inbreeding and LMC was made possible by the existence both of variation within species in the number of foundresses per fig (supplying variation in LMC) and of variation between species in the average number of foundresses per fig (supplying variation in inbreeding levels).

A paper by Yamaguchi¹ on page 460 of this issue also combines advanced theory and field data novel in at least two important respects. First, they add aphids to the list of major animal groups in which LMC appears to cause large sex-ratio biases. This is exciting news, in part because aphids have diploid genetics and chromosomal sex determination. How, then, do they control their sex ratios? The generation of aphids that infest plants throughout the summer are parthenogenic clones of XX females. But at the end of the summer, the last parthenogenic generation produces sexual females (who are XX) and males (who are XO). The male embryos are created by elimination of one X chromosome during a modified mitosis.

A second and strikingly novel pattern seen in Yamaguchi's data concerns sex-ratio differences among females who produce the sexual generation. These found-

Long-running experiments (II)



Although it has sometimes been stopped by mechanical failure, by the need for cleaning or when the ambient temperature has not fluctuated sufficiently, in principle this atmospheric clock has not needed winding since it was built in 1864. Now in the Physics Department of the University of Otago in Dunedin, New Zealand, its mechanism is driven by fluctuations in ambient temperature which expand and contract the air in an air-tight box. It is described in *European Journal of Physics* 5, 195-197; 1984, with a request for other examples of such experiments.

resses differ enormously in total fecundity yet, rather than producing the same proportion of males, each makes approximately the same absolute number. We are not convinced that Yamaguchi's simple and original theoretical explanation for this pattern captures all of the relevant features of the biology of the aphids, as it seems to assume that a female can assess the total fecundity of the foundress group to which she belongs; this would be reasonable for fig wasps, which produce their offspring after arriving at the fig, but may not be reasonable for the aphids studied by Yamaguchi, which produce their offspring before they fly to the trees on which the foundress groups are formed. But the pattern is there. It invites more theorizing and, of course, more empirical work.

XX-XO systems occur in many groups of arthropods and seem often to be associated with complex sex-allocation strategies (as in the laboratory nematode

Caenorhabditis elegans, which is mainly hermaphroditic, but with a small proportion of males). Yamaguchi's data should stimulate a widened search for instances of extreme sex-ratio manipulation among XX-XO species. Aphids themselves are economically important, so it may prove relatively easy to obtain support for studying their sex lives. □

1. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (Murray, London, 1871; reprinted by Princeton University Press, 1981).
2. Herre, E. A. *Science* **228**, 896 (1985).
3. Yamaguchi, Y. *Nature* **318**, 460 (1985).
4. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930; reprinted by Dover, New York, 1958).
5. Hamilton, W. D. *Science* **156**, 477 (1967).
6. Bull, J. J. *Evolution of Sex Determining Mechanisms* (Benjamin/Cummings, Menlo Park, 1983).
7. Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
8. Trivers, R. *Social Evolution* (Benjamin/Cummings, Menlo Park, 1985).

Robert M. May and Jon Seger are in the Department of Biology at Princeton University, Princeton, New Jersey 08544, USA.

Quantum Hall effect

Dimensionality leaves its mark

from R.J. Haug

WHAT happens when one plays with the dimensionality of an object? This is not an esoteric question of interest only to mathematicians but a very important problem that has been actively explored by solid-state physicists in the past few years, not least because it is likely to lead to many totally new physical phenomena. Already, measurements of a two-dimensional electron gas (2DEG) in strong magnetic fields have revealed the quantum Hall effect (von Klitzing, K., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45**, 494; 1980), for which von Klitzing was awarded this year's Nobel Prize in Physics. Now, H.Z. Zheng, K.K. Choi, D.C. Tsui and G. Weimann (*Phys. Rev. Lett.* **55**, 1144; 1985) have found a new, but related, effect — an anomalous variation in the magneto-transport coefficients of a two-dimensional electron gas as the width of the conducting channel becomes smaller than a characteristic length. This is of the order of $10\ \mu\text{m}$ depending on the mobility of the electrons. The authors assume that the effect, which was only observed in high-quality samples, arises from a transition from two-dimensional to one-dimensional behaviour.

One of the reasons for this branch of research is the need for ever smaller components in the electronics industry. But, in addition, the modulation-doped GaAs/AlGaAs heterostructures used by Zheng *et al.* can also be used to fabricate very fast switches (Störmer, H.L., *Nature News and Views* **317**, 20; 1985). One obtains these structures by growing thin layers of GaAs, AlGaAs and AlGaAs doped with silicon on a crystal of GaAs by molecular beam epitaxy (MBE). Because of the

different electronic properties of the two semiconducting materials GaAs and AlGaAs, the electrons from the Si donors in the AlGaAs transfer to the GaAs layer, so forming a quasi-two-dimensional electron gas bound to the interface via the Coulomb attraction of the ionized donors across the junction.

An electron localized in a region of sufficiently small thickness behaves differently from a free electron that is able to move over large distances in all directions. If the electron is confined within a narrow potential well in the third direction, the quantized energy levels become discrete rather than being quasi-continuous. The electrons are unable to move freely in this direction. At low enough temperatures and sufficiently small carrier concentrations all electrons are in the lowest energy level; this is a description of a 2DEG. Owing to the spatial separation of electrons from their ionized parent impurities, the scattering rate is strongly reduced, which leads to exceedingly high mobilities, particularly in the low-temperature regime.

If a strong magnetic field is applied in the z direction perpendicular to a 2DEG kept at the temperature of liquid helium, and the voltages in the direction of the applied current (V_x) and perpendicular to this direction and to the magnetic field (V_y) are measured, a non-vanishing V_y is found in the magnetic field. This is the effect found in 1879 by E.H. Hall at Johns Hopkins University and it is true for all conducting materials. Von Klitzing observed that in a 2DEG at certain regions V_y vanishes, although there is still a current I through the sample and the Hall

voltage V_x shows a plateau. In such regions the relation of V_y to the current I is given by $V_y/I = h/\eta e^2$, where h is the Planck constant, e the charge of an electron and η an integral number.

Zheng *et al.*, like von Klitzing, measured V_x and V_y as a function of the magnetic field at low temperatures, with the aim of investigating size effects in the sample by changing the width of the conducting channel. First, they noticed that the quantum Hall effect is exact. When V_y vanished, the Hall voltage always showed a plateau with $V_y/I = h/\eta e^2$. This is another proof that the quantum Hall effect is a universal property of two-dimensionality of the charge carriers and does not depend on properties of the samples, such as their size.

But Zheng *et al.* also looked at the lineshape of their experimental curves in the region between the plateaus of the quantum Hall effect. When the electrons had low mobilities no difference was observed between wide and narrow samples. But for samples with very high electronic mobilities, the lineshape changed dramatically when the width of the channel was reduced, although the measurements with a wide conducting channel were still comparable to those obtained with low electronic mobility samples.

The data suggest that the new phenomenon arises as the size of the sample approaches some characteristic length of the electronic system. One possible explanation put forward by Zheng *et al.* is that the characteristic length is that of the inhomogeneities in the sample. Another possibility is that it is associated with the electron-electron interaction, in which case the observed effect would have to be attributed to a dimensional crossover from two-dimensional to one-dimensional behaviour. In other words, if the channel is narrow enough the electrons cannot move freely in two dimensions: one direction is open but in the other motion is restricted by the edges of the sample.

These explanations are, however, not definitive. In experiments with silicon metal-oxide semiconductor field-effect transistors (Si-MOSFET), where samples of different channel length but identical channel width have been measured (von Klitzing, K. *et al. Proc. 17th Int. Conf. Physics Semiconductors*, San Francisco, 1984), very long samples produce results that are strikingly similar to those of Zheng *et al.* for very narrow GaAs/AlGaAs samples, whereas short samples do not. For a long sample it is difficult to explain the effect by a dimensional crossover as motion is not restricted in one direction. Nevertheless, the experiments give hints about the underlying fundamental physics. Further experiments and ideas will clarify these problems. □

R.J. Haug is at the Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1, D-7000 Stuttgart 80, FRG.

Cell biology

Organizing the nucleolus

from John Sommerville

THE function of the nucleolus is to produce ribosomes in response to the physiological demands of the cell. Nucleoli develop from chromosomal sites, known as nucleolar organizing regions (NORs), which contain tandemly arranged genes for ribosomal RNA (rRNA). The initiation and extent of their development is subject to both cell-cycle and metabolic influences. At a recent meeting* substantial advances were reported in identifying the components involved in nucleolar activity and in describing the interactions that occur between them.

After mitosis in a cell actively synthesizing proteins, the nucleolus is generated from the NORs as a massive and complex structure, bringing together all of the components required for the manufacture of ribosomes (see the figure). As well as transcription of the rRNA genes by RNA polymerase I under the control of associated transcription factors, this process involves many other components produced in other parts of the cell and brought into the nucleolus. These include about 80 ribosomal proteins, small RNA species required for rRNA processing (U3 and U8 small nuclear RNAs) and ribosome maturation (5S RNA), and several acidic proteins required for the maintenance of nucleolar structure and the assembly and transport of preribosomes. As these components have to be manufactured at a rate determined by the demands of ribosome production, considerable co-regulation of gene transcription is required.

Ultrastructural analysis combined with autoradiography to detect the initial sites of RNA synthesis (G. Goessens, University of Liège) and with selective silver staining to detect DNA- and RNA-associated proteins (K. Smetana, Czechoslovak Academy of Sciences, Prague) has made it possible to map the stages of ribosome formation to distinct regions in the nucleolus. The processes of initiation, production and maturation seem to progress from centre to periphery and the consensus is that the rRNA genes of the NORs are located in regions of low electron density (fibrillar centres), that gene transcription is initiated here but progresses into the surrounding dense fibrillar component where proteins bind to the transcripts, and that processing and maturation of pre-ribosomal particles occur in the cortical granular component.

A key event in initiating nucleolar development at the NOR is the activation of rRNA genes. It has been shown from immunostaining studies (U. Scheer, DKFZ Heidelberg) that substantial amounts of RNA polymerase I are located at NORs during mitosis, permitting rapid activation

of the genes as cells enter interphase. From *in vitro* reconstruction experiments using cloned genes and cellular extracts, the sequence of events in initiating the transcription of mammalian rRNA genes has been established (M. Muramatsu, University of Tokyo). First, a species-specific DNA-binding protein — transcription factor ID (TFID) — binds to the gene-promoter region located just before the transcription start site. A second, growth-dependent transcription factor (TFIA) then acts through protein-protein interactions to bring RNA polymerase I to the promoter. After this initiation complex is formed, both TFID and TFIA remain bound to the promoter region while RNA polymerase I reinitiates many rounds of transcription. It is interesting to note that point mutations at 7, 16 and 25 nucleotides before the transcription start site reduce activity to 0, 10 and 50 per cent of wild type, respectively. This nine-nucleotide interval could represent a linear series of protein contact points along the DNA double helix.

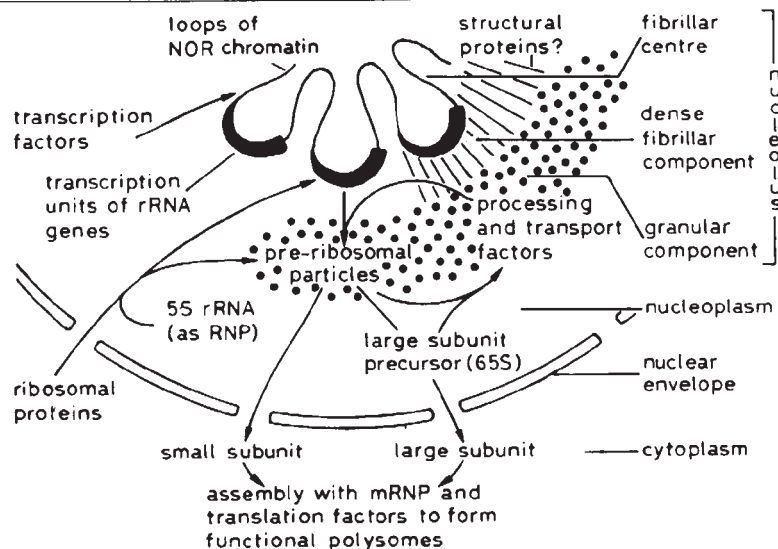
Attraction of transcription factors to promoters in *Xenopus* seems to take a special route (R. Reeder, Hutchinson Cancer Research Center, Seattle). In the spacer regions between rRNA genes, there are multiple 60 or 80 base-pair (bp) enhancer elements, a block of which confers a 20-fold increase in transcription activity compared with genes lacking enhancers. In gene constructs, these elements are effective when they are located 4–5 kbp from the gene promoter; they can be in either orientation and can be located in the gene itself. They seem to function as sites for the attraction and

entry of a transcription factor that is transferred to the gene promoter to form the transcription initiation complex.

Transcription of rRNA genes in *Xenopus* terminates 235 nucleotides beyond the end of the gene (Reeder). Using short DNA probes from this region, a stable RNA species corresponding to this extra transcription is detected, which tallies with electron microscope observations (M. Trendelenburg, DKFZ Heidelberg) of three additional polymerase molecules at the end of the rRNA transcription unit, with no long transcripts attached. It would appear that the rRNA transcripts are rapidly processed by cleavage before transcription terminates. Discovery that RNA polymerase I uses enhancer sequences and 3'-processing of transcripts brings its activities more into line with those of RNA polymerase II.

Of the many proteins that associate with rRNA gene transcripts at various stages of their processing, the best characterized are the ribosomal proteins (rps). Features contributing to the co-regulation of nucleolar rRNA genes and the many *rp* genes scattered throughout the genome have been studied in yeast (W. Mager, Vrije University, Amsterdam), *Xenopus* (E. Beccari, University of Rome) and mouse (K. Dudov, Bulgarian Academy of Sciences, Sofia). The *rp* genes of these organisms are similar in that they often lack a conventional TATA promoter sequence, having instead a long pyrimidine tract running through the transcription start site. In most of 22 *rp* genes in yeast that have been sequenced completely, there are two 12-bp-long homologous sequences implicated in transcription activation, located 300–500 nucleotides upstream of the initiation site.

Co-regulation of the primary rRNA-binding *rp*, L25 has been studied in detail by Mager. By increasing the dosage of L25 genes in yeast cells using a shuttle vector, it is found that L25 transcription is not



Steps in ribosome production in the nucleus

co-regulated with endogenous *rp* genes; instead excess L25 is synthesized, which in turn may block further translation of its own message. A link between nucleolar activity and L25 gene expression is provided by a homologous sequence present in 26S rRNA and L25 messenger RNA, either of which can bind the L25 protein.

The yeast system has similarities with the regulation of expression of *rpL1* in *Xenopus* (Beccari). Accumulation of L1 resulting from an increase in *L1* gene dosage by microinjection into oocytes, leads to blocking of the removal of introns from the *L1* gene transcript. A regulatory sequence is again implicated, here located in the introns of the *L1* gene transcript and in 28S rRNA. Binding of excess L1 to its own gene transcript would effect autogenous regulation by blocking processing and hence further production of L1 until it is removed by binding to rRNA.

In the nucleolus, the predominant proteins are neither transcription factors nor individual rps. Of particular interest are the silver-staining phosphoproteins with relative molecular masses around 100K. Their amino-acid sequences, derived from protein chemistry (H. Busch, Baylor College, Houston) and cDNA sequencing (M. Caizergues-Ferrer, CNRS, Toulouse), contain conserved functional domains. One, rich in glutamic and aspartic acids, is thought to interact with chromatin; another, rich in glycine and dimethylarginine, is found in various RNA-binding proteins. Because they are only transiently associated with rRNA gene transcripts and are lo-

cated in the inner fibrillar regions of the nucleolus, these proteins seem to be involved in relatively early stages of rRNA processing.

A completely different acidic nucleolar protein of 40K (B. Hügler, DKFZ, Heidelberg), associated with preribosomal particles in the cortical granular component of nucleoli and with the nucleoplasmic 65S precursor of the ribosomal large subunit, seems to be functional later. As it is never found outside the nucleus, its primary function seems to be the transport of preribosomal particles. Another important difference is that the silver-staining proteins, together with RNA polymerase I, remain associated with the NOR during mitosis whereas the 40K protein is distributed with some ribosomal protein over the entire surface of the chromosomes. Other nucleolar proteins are distributed into the cytoplasm during mitosis.

Obviously the mechanisms involved in bringing these dispersed components together to construct new nucleoli require much further study. One way into the problem might be to create novel NORs by microinjecting cells with recombinant plasmids containing rRNA genes that would fuse and integrate into chromosomal sites. This approach might also tell us whether intranuclear positioning is important — does the NOR have to be brought into close association with the nuclear envelope before nucleolar development can proceed? □

John Sommerville is in the Department of Zoology, The University, St Andrews, Fife KY16 9TS, UK.

100 years ago

The fresh-water mussel closes its shell by contraction of two strong muscles — one before and one behind; but how does it open its shell? This has recently been studied by Herr Pawlow. The animal was fixed on a board by one shell, while the other shell was connected by a silk thread with the short arm of a lever, the longer arm of which indicated the movements on a slowly-rotating drum. . . . there are two classes of nerve-fibres connected with the muscles — the one motor, producing contractions; the other inhibitory, producing relaxation.

From *Nature* 33 106, 3 December 1885.

an event that took place in the late Pliocene, 2.3 Myr ago. A search in contemporaneous sediments recovered from the North Pacific sea floor has failed to detect an equivalent siderophile anomaly. Although the exact spatial distribution of the Antarctic debris is unknown, the event was clearly not global in scale, unlike the event thought to have occurred at the Cretaceous–Tertiary boundary.

The estimated infall of meteoritic material in one of the Antarctic cores is about 100 mg cm^{-2} . This rivals the most recent estimate of the average global outfall of meteoritic (chondritic) material of about 140 mg cm^{-2} at the Cretaceous–Tertiary boundary by F.T. Kyte, J. Smit and J.T. Wasson (*Earth planet. Sci. Lett.* 73, 183; 1985). In this core, enrichments in the siderophile element iridium are obviously correlated with the meteoritic material and reach a maximum of approximately 5 ng g^{-1} . This is roughly an order of magnitude less than the maximum values reported for the Cretaceous–Tertiary boundary at Caravaca, Spain. Variation in the relative abundances of siderophiles is a geochemical characteristic of various meteorites and the difference indicates that the Antarctic projectile differed in composition from that postulated to have caused the Cretaceous–Tertiary iridium anomaly.

Because the Antarctic debris is of relatively coarse-grained nature, with particles up to several millimetres in diameter, the type of projectile can be identified with some confidence. Three types of particle have been recovered: vesicular glasses; brecciated and shocked basalt; and Fe–Ni metal. Most of the iridium is in the vesicular particles, which make up more than 90 per cent of the recovered particles and contain about 150–200 ng iridium per gram. The vesicular particles have some textural similarities to meteor ablation spheres but their composition is similar to that of basalt. The identifiable basalt particles consist primarily of the minerals plagioclase and pyroxene and are often enclosed in vesicular material. The basaltic particles could originate either from projectile material or from ocean-floor basalts. Although no chemical analyses are presented, it is clear from the mineral chemistry, particularly the manganese content of the pyroxenes, that they represent material from the group of meteorites termed basaltic achondrites, probably of the howardite type.

Meteorites

Siderophile-enriched sediments and meteoritic debris

from Richard A. F. Grieve

Low sedimentation rates make deep-sea sediments a relatively rich depository for the extraterrestrial materials that bombard the Earth. F. T. Kyte and D. E. Brownlee (*Geochim. cosmochim. Acta* 49, 1095; 1985) give details of an important find of meteoritic debris in the Antarctic Basin, which provides evidence that sediments chemically enriched in siderophile elements, those elements that tend to concentrate with iron, are associated with meteoritic material. This discovery, initially reported in *Nature* (292, 417; 1981), clearly demonstrates that large impact events can produce siderophile-enriched sediments, the postulated explanation for the origin of the iridium anomaly found at the Cretaceous–Tertiary boundary.

Siderophile elements, which include

cobalt, nickel, iridium and platinum, occur preferentially with native iron. As a result, they are relatively depleted in the Earth's crust and concentrated in the core. Certain meteorites, however, are relatively rich in siderophiles, and local enrichments in siderophiles have been found in so-called impact melt rocks associated with some terrestrial meteorite impact structures. It had been suggested that siderophile-enriched sediments could also result from the presence of meteoritic material but until now no direct association of siderophile-enriched sediments and meteoritic material had been observed.

The meteoritic material was discovered in two piston cores, recovered from water approximately 5 km deep at sites 120 km apart, at 57°00'S, 89°12'W and 57°47'S, 90°48'W. Palaeomagnetic stratigraphy indicates that the meteoritic debris is from

*The Ninth Nucle(ol)ar Workshop, Krakow, Poland, 13–17 September 1985.

This poses a problem. Howardites are differentiated meteorites; that is, their source body underwent large-scale chemical differentiation with the formation of an iron-rich core and a silicate crust. Thus the howardites, which represent the brecciated silicate crust, are relatively depleted in siderophile elements. The basaltic clasts, therefore, cannot be the source of the siderophiles in the vesicular material. The most obvious source is the rare-metal grains. Calculations using a mixing model indicate that the vesicular material could correspond approximately to howarditic material with a 3 per cent metal component. The metal composition required by the mixing model does not correspond exactly to that determined for the cores. This is not, however, a major difficulty, as only one grain was analysed and changes in metal composition caused by oxidation reactions during impact have been documented at terrestrial impact sites.

If the vesicular glasses are a mixture of howarditic material and metal, with a minor addition of seawater, what was the thermal event that led to melting? Heating during atmospheric entry or impact with the ocean are two possibilities. Although they present no compelling arguments, the authors favour the latter. Impact as the heat source is reasonable, at least from the point of view of physics. If impact occurred at 15 km s^{-1} , peak shock pressures would be of the order of 150–200 Gpa, with the attendant post-shock temperature well above the melting temperature of howarditic materials. If the thermal energy

associated with impact was sufficient to produce the vesicular particles, then why are they intimately associated with unmelted basaltic particles? The authors suggest the projectile suffered at least partial break-up in the atmosphere. Smaller pieces would be decelerated by atmospheric drag, so that their impact velocity would be below that required for melting, and would recombine with the melted vesicular material through collisions in the cloud of ejecta. The size of the projectile is poorly constrained at approximately 100–500 m in diameter, on the basis of the distribution of meteoritic material. A projectile in this size range would be subjected to some atmospheric crushing and dispersion, according to the calculations of H.J. Melosh in *Multi-Ring Basins*, 29 (Pergamon, 1981).

From recent calculations of the terrestrial cratering rate, it is estimated that as many as 50 similar events occur in the Earth's oceans every million years. As ocean-floor sampling continues and deep-sea sediments are examined, it is likely that other oceanic impact events will be discovered. This Antarctic discovery not only strengthens the meteoritic origin of the siderophile enrichments at the Cretaceous–Tertiary boundary but, together with potential future discoveries should also provide insights into the complementary subjects of meteoritics and impact phenomena. □

Richard A.F. Grieve is in the Earth Physics Branch, Energy, Mines & Resources Canada, Ottawa K1A 0Y3, Canada.

pleted. They chose to examine the nuclear DNA content of plants growing in a north-facing area of damp limestone grassland in North Derbyshire, England, that was particularly rich in herbaceous plant species. They removed 150 circular turf samples of 10 cm diameter and maintained them under garden conditions, where they were able to establish the rates and times of leaf elongation for all the major species. At the same time they monitored temperature, and measured nuclear DNA content in samples of plants' root tips.

The nuclear DNA contents of the plants was fairly evenly distributed over a wide range from less than 1 pg to almost 40 pg per nucleus. In the cold conditions of early spring (March), species with more than 10 pg DNA per nucleus grew two to four times as fast as those with a less than 10 pg per nucleus. By the beginning of May, the growth rates of the two groups were only slightly different and by June the differences were entirely insignificant.

The experiment clearly demonstrates the limiting effect of temperature on leaf extension in the early part of the year and also the advantage gained in this period by species with a high nuclear DNA content. The authors regard a small nuclear DNA content as characteristic of plants which grow by rapid and continuous cell division during favourable periods (such as the summer season in temperate areas). Rapid spring growth is best achieved by expanding cells which were produced during the previous growing season but which have remained unexpanded during the intervening winter. The problem of conducting mitosis under cold conditions is thus avoided. Such plants characteristically have high nuclear DNA contents. Their growth declines in summer when they conduct repeated cell divisions to produce tissues which will expand during the following spring.

From an ecological point of view, one of the most interesting points to emerge from this work is the chance it offers to catalogue at least one dimension of niche specialization simply on the basis of cytological investigation of the plants in a community. An examination of the spread of DNA values in a community, for example, provides an insight into the extent that the various species are exploiting the range of temporal opportunities available to them.

Now that it has been demonstrated that the broad relationship between DNA content and the time of canopy expansion actually operates in a natural community, it should be possible to make informed predictions about the temporal role of species in various communities on the basis of their DNA content. As Grime and his colleagues point out, DNA analyses may also provide a clue to the vexed question of how certain vigorous community dominants can coexist. For example, the fact that one robust grass, *Brachypodium pinnatum*, has a nuclear DNA content of

Ecology

Nuclear DNA content as a guide to plant growth rate

from Peter D. Moore

ONE of the most clearly defined aspects of the ecological niche of a plant is the temporal element in its seasonal growth and development¹. Even if all the plants in a meadow are tapping the same energy resource, they can partition that resource in time by expanding their leaf canopies in sequence. A knowledge of such differences can be of considerable importance in the management and conservation of grasslands where decisions must be made concerning the timing of mowing or grazing regimes². As a result of the recent work of Grime, Shacklock and Band³, it is becoming apparent that the timing of the growth of the shoot systems of many grassland plants is closely related to their nuclear DNA content, which means that anatomical examination can provide considerable information about the nature of niche partitioning within a plant community.

Grime and Mowforth⁴ had already suggested that the nuclear DNA content of certain plants is related to the timing of

shoot growth. Having observed that Mediterranean species of plants, which have their main growth periods in winter and early spring, have a high nuclear DNA content, they examined a wide range of British plant species to see whether any relationship between DNA content and time of growth was apparent. The found that species like *Hyacinthoides non-scriptus* (bluebell) and *Ranunculus ficaria* (lesser celandine) that commence growth in the early spring (April) have DNA contents that are about ten times as high as those that commence growth rather late in the season (June), like *Urtica dioica* (stinging nettle) and *Molinia caerulea* (purple moor grass). The capacity to grow at lower temperatures is thus associated with the possession of a large genome.

But the evaluation of the full ecological implications of this proposal required the analysis of an area of vegetation under natural field conditions, which is what Grime and his co-workers have now com-

2 pg. whereas another. *Bromus erectus*, contains 22 pg per nucleus, may explain the way in which these species partition the environmental resources and thereby manage to coexist. Studies of this kind will assist botanists to appreciate more fully the implications of Gause's competitive exclusion hypothesis in the realm of vegetation. □

1. Al-Mufti, M.M., Sydes, C.L., Furness, S.B., Grime, J.P. & Band, S.R. *J. Ecol.* **65**, 759 (1977).
2. Duffey, E. *Grassland Ecology and Wildlife Management* (Chapman and Hall, London, 1974).
3. Grime, J.P., Shacklock, J.M.L. & Band, S.R. *New Phytol.* **100**, 435 (1985).
4. Grime, J.P. & Mowforth, M.A. *Nature* **299**, 151 (1982).

Peter D. Moore is in the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF, UK.

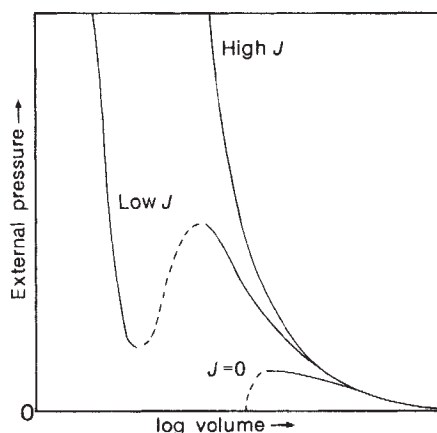
Astrophysics

Phase transitions in star formation

from Alan P. Boss

MANY dynamical aspects of star formation have become understood through theoretical models of collapsing interstellar clouds but attention has recently been turned to equilibrium models. From models of the possible equilibrium states available to interstellar clouds and their protostellar progeny, it emerges that two types of equilibrium may occur for isothermal clouds with small angular momentum, with one being much denser than the other. Therefore, if there are phase transitions — mathematically analogous to the transition of a Van der Waals gas to a liquid — from the more diffuse to the more compact state, the result could be star formation in clouds that would otherwise be considered stable¹.

An interstellar cloud must contract and increase its density by a factor of about 10^{20} in order to form a star. Observation has yielded little knowledge about the contraction process, because of the inherent obscuration and small length scales involved. Theoretical models of nonmagnetic star formation have focused on solving the exceedingly difficult time-dependent problem of the collapse of an interstellar cloud from an initially unstable state.



Axisymmetrical equilibrium sequences for isothermal clouds with differing amounts of angular momentum, J . No equilibrium exists for sufficiently dense clouds with $J = 0$. Clouds with low J can exist in either a diffuse or compact state, separated by unstable equilibria (broken line). For high J clouds the unstable branch disappears. Phase transitions from the diffuse to the compact branch may initiate star formation (modified from ref. 2).

To include rotation and the possibility of fragmentation, the model must include all three spatial dimensions, a formidable task. Protostellar clouds are effectively isothermal during much of their evolution, but eventually they become opaque and heat up, adding further to the complexity of the computations.

A simpler approach is to seek equilibrium states for rotating, isothermal, axisymmetrical, self-gravitating clouds (see figure). Models based on the global properties of spheroids of uniform density reveal that, for clouds with low angular momentum, two types of equilibrium are possible²⁻⁴: slowly-rotating, nearly-spherical clouds, supported primarily by external pressure (diffuse branch), and rapidly rotating, highly flattened clouds, supported largely by rotation (compact branch). Intermediate configurations are unstable to compression. For clouds with large angular momentum, a single sequence exists, extending monotonically to high densities.

The diffuse state is metastable, because its free energy is higher than the compact branch. The unstable branch has the highest free energy, forming a barrier between the stable branches. If the free energy of a diffuse branch cloud is raised sufficiently, such as by compression in the shock front associated with the spiral arms of our Galaxy, an otherwise stable cloud could dynamically become a compact cloud. Such a phase transition would allow low angular momentum clouds that are stable against collapse to become significantly denser¹.

Phase transitions can only be considered physically possible if detailed models of the two branches can be constructed and shown to be dynamically stable. Rigorous modelling of clouds with external pressure produces stable clouds that become progressively more condensed at their centres as the mass increases⁵. The models become unstable to compression at high central density, as in non-rotating, isothermal clouds supported by external pressure. These models are the realistic manifestation of the diffuse branch³.

Analytical solutions exist^{6,7}. They are compressionally unstable when the cloud is nearly spherical, and dynamically un-

stable to ring formation when highly flattened⁸; intermediate models are stable to axisymmetrical perturbations. However, these solutions are mathematically singular at the origin and extend to infinity. Because of the lack of external pressure, the flattened solutions are similar to the compact branch. Thin-disk models have been constructed which also suffer from an infinite central density, but they do include external pressure⁹. These models are the closest yet developed to a rigorous description of the compact branch. Because compact branch models with finite central density have not yet been constructed, and their stability to axisymmetrical perturbations has not been tested, it is premature to declare the compact branch a possible member of the physical universe. It is important to note that calculations of the dynamical collapse of isothermal clouds have found the diffuse branch¹⁰ but not the compact branch. Instead, collapsing axisymmetrical clouds produce either rings or runaway disks, depending on details such as the initial distribution of angular momentum.

This discussion has dealt with stability to compressional or axisymmetrical perturbations but interstellar clouds can also distort in a non-axisymmetrical way. Simplified ellipsoidal models form a high angular momentum sequence (similar to the diffuse branch) that terminates when sufficiently compressed¹. The simplified compact branch models are usually dynamically unstable to non-axisymmetrical distortions which could result in fragmentation and further contraction to the stellar state¹. The stability of detailed models of the compact state⁹ to non-axisymmetrical distortion is unknown.

If phase transitions occur between diffuse and compact interstellar clouds, it may be easier than has been supposed for the process of star formation to be initiated. Clouds that are stable to small perturbations may be hammered into a dynamic collapse towards a denser configuration. A collapsing cloud may well bypass any compact equilibrium state and fragment directly into a protostellar system. The validity of these speculations awaits confirmation of the existence and stability of the compact branch. Bear in mind, however, that unstable equilibrium states are of considerably less interest to nature than to theoretical astrophysicists. □

1. Tohline, J.E. *Astrophys. J.* **292**, 181 (1985).
2. Weber, S.V. *Astrophys. J.* **208**, 113 (1976).
3. Hachisu, I. & Eriguchi, Y. *Astron. Astrophys.* **140**, 259 (1984).
4. Tohline, J.E. *Icarus* **61**, 10 (1985).
5. Stahler, S.W. *Astrophys. J.* **268**, 165 (1983).
6. Hayashi, C. *et al. Prog. theor. Phys.* **68**, 1949 (1982).
7. Toomre, A. *Astrophys. J.* **259**, 535 (1982).
8. Schmitz, F. *Astron. Astrophys.* **131**, 309 (1984).
9. Hachisu, I. & Eriguchi, Y. *Astron. Astrophys.* **143**, 355 (1985).
10. Boss, A.P. & Haber, J.G. *Astrophys. J.* **255**, 240 (1982).

Alan P. Boss is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington DC 20015, USA.

New strategies for AIDS therapy and prophylaxis

SIR—In search for a reactive agent against HTLV-III/LAV, the virus that causes AIDS (acquired immune deficiency syndrome), the attention is focused on neutralizing antibodies. Such antibodies have been detected in the blood of AIDS patients^{1,2} and form the starting point for the production of vaccines³. However, there are as yet no indications that these antibodies influence the course or the outcome of the disease. Therefore, alternative strategies should be pursued. A recently developed technique for specific inhibition of gene expression provides one possible new way to tackle the AIDS virus. It is based on the inhibition of the messenger RNA translation through the interaction of anti-sense RNA with a specific messenger and in several systems has been shown to be very successful⁴⁻⁶. Accordingly, the introduction of an actively transcribed template for anti-sense RNA of one of the AIDS virus genes into infected cells may induce inhibition of virus replication and increase survival chances of the patient.

The construction of such a template will not be difficult, but how is it to be delivered to the correct target cells? The answer is, in principle, simple. Make a recombinant AIDS virus in which one of the genes is in the reversed orientation. Infection of a patient with this virus will ensure that the recombinant genome is delivered precisely to those cells that are sensitive to, or harbour, the AIDS virus. Cells not yet invaded by the AIDS virus may be protected by the recombinant anti-sense RNA against integration of the AIDS genome. In this way spreading of the virus within the body could be prevented. Which segment of the genome of the virus should be reversed? The *gag* or *pol* regions are unattractive, because the corresponding genes of any helper virus involved in the production of the recombinant might be blocked as well. With the reversion of the *env* gene there is the additional problem that the recombinant virus might acquire the envelope of the helper, resulting in the infection of the wrong target cells. A fourth gene, *tat*, which has recently been identified on the genome of the AIDS virus^{7,8}, is more promising. The gene product is a trans-activator of viral transcription. Preventing the synthesis of the *tat* protein results in a significant reduction of virus replication, which is exactly what we are looking for. In fact, reversion of the second exon of the *tat* gene, which is located between *pol* and *env*, would alone be sufficient.

In theory, the only problem that we are left with is the efficient production of the recombinant virus, because this may require the action of the *trans*-activator protein. A cell line producing and accumulating the *tat* protein in sufficient quantity, prior to infection with the recombinant

virus, might be effective. Now that the sequence of the gene is known such a cell line is within reach. Alternative methods can be thought of, one of which is to load cells before infection with the *tat*-protein, obtained from any source, by the method of scrape loading⁹, which works well especially with proteins of small size.

Although problems would surely show up during the development of the strategy outlined above, AIDS is too serious a threat to human life to let any possible means for therapy and prophylaxis go untested. The method might also be tried for other diseases, including the T-cell leukaemias caused by HTLV-I and -II.

E.C.M. MARIMAN

Department of Human Genetics,
University of Nijmegen,
6525 GA Nijmegen,
The Netherlands

1. Weiss, R. A. *et al.* *Nature* **316**, 69 (1985).
2. Robert-Guroff, M. *Nature* **316**, 72 (1985).
3. Crowl, R. *et al.* *Cell* **41**, 979 (1985).
4. Izant, J. G. & Weintraub, H. *Science* **229**, 345 (1985).
5. Melton, D. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 144 (1985).
6. Rosenberg, U. B. *et al.* *Nature* **313**, 703 (1985).
7. Arya, S. K. *et al.* *Science* **229**, 69 (1985).
8. Sodroski, J. *et al.* *Science* **229**, 74 (1985).
9. McNeil, P. L. & Taylor, D. L. *Cell Calcium* **6**, 83 (1985).

SIR—We wish to suggest a gene therapy scheme for the control of acquired immune deficiency syndrome (AIDS). The current view on AIDS is that infection with HTLV-III/LAV virus, trophic for OKT4⁺ lymphocytes, results in breaking the backbone of the immune system. The diversity of viral strains and the nature of the infection renders problematic the production and use of vaccines¹. A strategy has been proposed consisting of "protecting" T-cell clones against HTLV-III by using viral interference through infection with attenuated virus².

In our view a better way to achieve this protection would be through the strategy of antisense RNA. Inhibition of gene expression has been achieved in a variety of systems with as little as 52-base-pair, homology to the 5' untranslated segment of the sense RNA³. Helper-free, defective HTLV-III, containing the appropriate group-specific inverted viral gene segments could be produced in a packaging cell line modelled on the Zipneo-ψ2 system⁴. To ensure out-titration of sense-RNA, a gene amplification system, such as the DHFR gene responsive to the anticancer drug methotrexate, could be built into this antisense virus⁵.

AIDS therapy would consist in removal of bone marrow stem cells from patients, followed by their infection with the anti-virus and reinsertion, resulting in the production of "protected" lymphocytes. Alternatively, peripheral blood lymphocytes obtained by pheresis would be infected and reinserted. Methotrexate induced amplification of the antisense provirus may be carried out in vitro, as needed⁶. By these means a critical and sufficient population of T cells might be protected so that an effective immune

response against all strains of HTLV-III and against opportunistic infections could be achieved.

RAYMOND TELLIER

JOSEPH M. WEBER

Department of Microbiology,
Centre Hospitalier Universitaire
de Sherbrooke,
Québec, Canada J1H 5N4

1. Wong-Staal *et al.* *Science* **229**, 759 (1985).
2. Bolognesi, D. P. & Fischinger, P. J. *Cancer Res.* **45**, 4700s (1985).
3. Izant, J. G. & Weintraub, H. *Science* **229**, 345 (1985).
4. Cepko, C. L. *et al.* *Cell* **37**, 1053 (1984).
5. Miller, A. D. *et al.* *Mol. Cell. Biol.* **5**, 431 (1985).
6. Schornagel, J. H. & McVie, J. G. *Cancer Treat. Rev.* **10**, 53 (1983).

Cerebral supermanifolds and little brains

SIR—John Marshall's analysis of current neuropsychological efforts to correlate specific lexical dysfunctions with "focal" brain damage raises, expectedly, some presently unfocused issues: (1) The "focal" brain damage reported by Hart *et al.*¹ was a "left-hemisphere cerebrovascular accident" diagnosed by computerized tomography as "an infarction involving the left frontal-lobe and basal ganglia." This is, by cerebral neuroanatomical criteria, a truly enormous, almost global region of damage.

(2) Marshall's interesting quotation from Johann Gesner² suggests that even the classical eighteenth century allowed speculation about the validity of linear transformations from a 'psycho'-logical to a cartesian frame of reference in cerebral cortex.

(3) Despite some rough approximations to a two-dimensional somatotopy found in mammalian and primate somatosensory-specific cortex (for example, the primate digit and mouse whisker-barrel "representations"), we have no reason to suppose a linear two- or three-dimensional mapping of inner and outer spacetime to exist inside our heads. Since the days of Maxwell to Einstein would it not be more reasonable to use differential geometry in our approach to understanding how brains might work? A start has already been made for our "little brains".

HILTON STOWELL

ERBP Laboratory,
120 Nature Creek SW,
Milledgeville Georgia 31061, USA

1. Hart, J., Berndt, R. S. & Caramazza, A. *Nature* **316**, 439 (1985).
2. Marshall, J. C. *Nature* **316**, 388 (1985).
3. Pellionisz, A. & Llinás, R. *Neuroscience* (1977-85).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a rather technical character which are not provoked by articles or letters previously published (where the Matters Arising section remains appropriate).

Feathered wisdom

Freeman J. Dyson

Hawks, Doves and Owls: An Agenda for Avoiding Nuclear War.

Edited by Graham T. Allison, Albert Carnesale and Joseph S. Nye, Jr.
W. W. Norton: 1985. Pp. 282. \$14.95, £12.95.

THIS book has been edited by three pillars of the Harvard establishment who have long experience as expert advisors to the American government. They have produced an introductory chapter and two summarizing chapters. In between are six chapters written by others, discussing in greater depth six particular contexts within which nuclear war might arise. Each is a scholarly analysis, full of historical facts and supported by numerous footnotes. The book as a whole is readable and tells a coherent story without becoming repetitious. Its main emphasis is put on confusion as the most likely cause of nuclear war.

The longest chapter is also the most informative. It is "Soviet Perspectives on the Paths to Nuclear War," by Stephen Meyer, director of the Soviet Security Studies Working Group at MIT. Unusually, for an American discussing strategic questions, Meyer has paid careful attention to what the Soviet military authorities say and do. Half his extensive footnotes refer to Russian-language sources. But more than this he says: "Let me begin, then, by noting some sources not used in this chapter. First in this category are official Soviet pronouncements on military doctrine and policy . . . A second source not used is the work of Soviet academics . . . This chapter's sources include Soviet military planning literature, Soviet military historical analyses, Soviet diplomatic behavior, Soviet military force structure and deployment data, and memoir material." Most of us amateurs who claim to know something about Soviet attitudes collect our impressions from the two sources which Meyer discards as unreliable. His conclusions are firmly based and not proffered where evidence is lacking.

Meyer's description of Soviet attitudes emphasizes the Soviet doctrine of pre-emption. He does not overlook the inconsistencies between the formal 'no-first-use' policy enunciated by Brezhnev and the practical reliance on pre-emption which is visible in Soviet operational planning. The inconsistencies will not be quickly or easily resolved. We see a clear recognition by the Soviet military that decisions on the use of nuclear weapons are restricted to the supreme political authorities, and we see military plans which take the political decisions for granted.

Meyer believes that in an acute crisis the military plans are likely to prevail. The

most serious danger of nuclear war would arise from a Soviet decision to launch a pre-emptive attack in response to a sudden and panicky mobilization of NATO forces at a moment of crisis. He regards as particularly unwise the American habit of putting nuclear forces on world-wide alert as a signal of political resolve. This was done during the Cuban crisis of 1962 and the Egyptian crisis of 1973, in both cases without causing disaster. If it is done again in a future crisis arising closer to Soviet territory and appearing to threaten Soviet vital interests, disaster may be more difficult to avert.

Another good chapter is "Escalation in Europe" by Fen Osler Hampson, a Harvard post-doctoral fellow. One of Hampson's important suggestions is that, in case of serious fighting in Europe or in the Middle East, the escalation to nuclear war may happen more easily at sea than on land. Naval units on both sides have a freer access to nuclear weapons than land units, and the political controls at sea are looser. At sea there is an abundance of tempting targets for nuclear weapons, and the absence of civilian populations from the battlefield makes it easier for the warriors to yield to temptation. For all these reasons, the nuclear weapons at sea are more likely than those on land to get us into an unintended nuclear war.

There is an outstanding opportunity for us here to lessen the risk of nuclear war by extending and reinforcing the American-Soviet Incidents-at-Sea Agreement of 1972. The Incidents-at-Sea Agreement has two main components. It formulated an agreed set of rules for regulating traffic

in areas such as the Mediterranean where the two navies are apt to encounter each other frequently. And it established a joint commission of American and Soviet naval officers who meet twice a year to discuss dangerous incidents and to revise the rules when necessary. The holding of regular and quiet meetings between military officers is perhaps the most useful immediate step we can take to decrease the probability that local crises will escalate to general war. There is no reason why similar meetings should not be held by armies and air-forces as well as by navies.

The final summary by the editors is divided into two parts. First they give us an analytical framework for thinking about nuclear strategy; then they tell us what to do. The framework is based on the three ways in which people are afraid that a nuclear war might begin. War might begin because national leaders are weak and irresolute in standing up to aggression, or because they are paranoid in reacting to imagined or exaggerated threats, or because they are confused and unable to control events at moments of international turmoil. People who are most afraid of weakness are hawks. People who are most afraid of paranoia are doves. People who are most afraid of confusion are owls. The editors remark that all three fears are well justified by the evidence of past history and present realities. Hawks, doves and owls are all speaking truth.

A rational strategy must deal with all three fears simultaneously. What to do, is a question of balance and emphasis between the three dimensions of danger. The debate until now has been largely between hawks and doves, and has for that reason become ideological and impractical. The editors seek to redress the balance by giving major emphasis to the concerns of the owls. One of the wisest of the owls is William Ury, whose recent book *Beyond the Hot Line* gave us a number of sensible suggestions to mitigate the dangers of war arising from accident, incompetence and panic. Although Ury is not one of the present authors, his ideas are plainly visi-

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Popperfoto

High tension — the USS Barry steams alongside the Soviet freighter Anosov off Puerto Rico during the Cuba crisis of 1962. Could we now cope better in such a situation?

ble in many of the chapters. He was a member of the Avoiding Nuclear War Project out of which the book grew.

The final list of recommendations is divided into ten sections. Two of the sections are hawkish, two are dovish and six are owlish. These proportions reflect the editors' belief that at present the dangers of war through weakness or paranoia are over-rated and the dangers of war through muddle are under-rated. The majority of their recommendations are directed toward avoiding a nuclear version of the outbreak of war in 1914, an outbreak driven by geographical and historical accident rather than by any grand design.

A small sample may indicate the general flavour of the editors' recommendations. I choose five that I consider characteristic, one hawkish, "Don't adopt a No-First-Use policy," one dovish, "Don't assume that nuclear deterrence will last forever," and three owlish, "Do reduce

reliance on short-range theater nuclear weapons, do work with the Soviets to prevent and manage crisis, do encourage non-governmental contact with the Soviets." It is not necessary to agree entirely in order to applaud the clarity of their thinking and the brevity of their style.

I happen to disagree strongly with their No-First-Use recommendation, and still I recommend their book whole-heartedly as a fair and lucid statement of the nuclear dilemma. Only one essential thing is absent from their world-view. They lack a sense of the absurd. It may be that a resolution of the nuclear dilemma will come in the end from a general recognition of its absurdity. I would like to add one more recommendation to their list, "Don't take nuclear weapons more seriously than they deserve." □

Freeman J. Dyson is in the Institute for Advanced Study, South Olden Lane, Princeton, New Jersey 08540, USA.

Cool side of life

Pierre Douzou

Biophysics and Biochemistry at Low Temperatures. By Felix Franks. Cambridge University Press: 1985. Pp. 210. £25, \$44.50.

ALTHOUGH Felix Franks in a sense is correct in his forward "cold is the fiercest enemy of many forms of life", biologists dream of exploiting it for the cryopreservation of "spare parts" of living systems amongst other problems. Like physiologists some decades ago, biologists are becoming increasingly interested in the measurement, interpretation and exploitation of natural phenomena at low temperatures, and cryogenic research is slowly becoming a branch of this discipline. Franks' book is a step along the way in this endeavour.

The book gives clear answers to the classical questions about the physics and physical chemistry of water at sub-zero temperatures, the effects of its behaviour on life, and current applications in biochemistry. Other specialized chapters survey our understanding of the responses of the cell and of living organisms to cold, and of the problems of their preservation in the laboratory, a problem still dominated by empiricism and clouded by the interplay of an extraordinary variety of factors. The book ends with a survey of future prospects in this challenging but promising field.

Franks has been a leading contributor to research on water structure and properties, here he combines this expert insight with his interest and recent experience in cryopreservation for the task of explaining what is known and what is uncertain, and of suggesting what can be done to set up a useful cryobiology. The writing of a book on this topic demands a grasp of many specialized fields and is a daunting task, but the result here is agreeable and very instructive, because the book reveals its origin as a lecture course by a man trained as a physical scientist who drifted into the life sciences "fairly late in life and [who] never received any formal teaching in biological dogma" and had first to teach himself.

The book gives a strong impression of unity, carries with it a tone of optimism and in some cases of certitude and is valuable in that it covers a range of subject material which has never before been brought together in such an organized and concise yet clear manner. It will be of great interest to that growing band of researchers working on the cryobehaviour of living organisms. □

Pierre Douzou is Professor at the Museum d'Histoire Naturelle, Paris and head of Department at the Institut de Biologie Physico-chimique, 13 rue P.M. Curie, 75005 Paris, France.

Manual works

Graham Richards

A Handbook of Computational Chemistry: A Practical Guide to Chemical Structure and Energy Calculations. By Tim Clark. Wiley: 1985. Pp. 332. £35.80, \$35.

Semi-Empirical Methods of Quantum Chemistry. By Joanna Sadlej. Ellis Horwood: 1985. Pp. 386. £47.50, \$71.25.

COMPUTATIONAL chemistry has come out of the closet. No longer is the calculation of molecular conformational energy or electronic distribution confined to a small clique of tolerated but mildly deviant academics who scorn conventional experimental practice. Even in the staid halls of industry, theoreticians are now to be found working alongside straight organic chemists, designing novel compounds with the aid of calculations — both quantum mechanical and of the molecular mechanics type — and, increasingly, assisted by sophisticated computer graphics displays.

This growing army has a real need for a handbook which details not the theory behind the methods used, but rather where to start and how to use the program packages available. The most favoured weapons currently in use for elucidating conformation are programs based on molecular mechanics (treating molecules as balls and springs with empirical force fields). These are supplemented with molecular orbital calculations to unravel the electronic details, with the semi-empirical methods known by the acronymic titles MINDO, MNDO and MOPAC becoming standard, together with variants on the *ab initio* molecular orbital method developed by Pople and his co-workers (currently known as GAUSSIAN 82).

The use of all these packages requires

the chemist to present as input the geometrical structure of the molecule under consideration. Clark's excellent manual provides clear and detailed instructions on how to submit data for the most widely adopted packages. It is replete with examples of sample input and of output, just the thing to enable an experimental chemist to be able to use the available programs. The book does not provide much about the general background theory, nor does it underline the rather shaky theoretical basis for some of the methods (these matters are, however, discussed in many, more conventional textbooks on theoretical chemistry). Instead we have a very practical manual, the review copy of which has already been tested by novice graduate students with happy results. The theoretical section of chemistry bookshelves is generally overstocked, but this book fills a genuine gap and will undoubtedly be found at the elbow of many users of packages.

Sadlej's text is in an overlapping area and is a more conventional account of the semi-empirical methods of quantum chemistry: sadly it is the book one wanted ten years ago. The Polish edition was written in 1976 and although "revised", the age shows. The bibliography goes up only to 1978, and the techniques upon which the book concentrates — the Hückel method and CNDO — are both now somewhat *passé*.

Those theoretical chemists who resent the loss of their exclusive and even furtive world will no doubt grumble that running a computer package no more makes one a theoretician than sleeping in a garage makes one a mechanic. But computational chemistry is on the rise and life should be made easier for the virgins. □

Graham Richards is a lecturer in the Department of Physical Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK.

To win friends and influence people

Ian Stewart

The Elements of Graphing Data. By William S. Cleveland. Wadsworth: 1985. Pp. 323. Hbk \$27.95, £37.80; pbk \$18.95, £25.60.

"In 1980," says the author, "I began a study of graphs in scientific publications . . . In one study, I read the articles and reports of the 1980 Volume 207 of *Science* . . . 30% of the graphs in the volume had at least one of four types of specific problems." Those problems were: something on the graph but not explained, parts of the graph being visually indistinguishable, and mistakes such as mislabelling or poor reproduction. Deeper problems included poor choice of the form in which to present the data, and indeed poor choice of data. In an effort to combat these problems, and to suggest effective counter-measures, this book discusses the principles and methods of graphical construction, amply illustrated by examples selected from the literature (with references, which may lose the author a few friends but makes the arguments more cogent).

Scientists spend a large part of their training learning how to do science; and very little learning how to communicate it. Perhaps this is unavoidable, especially given the pressures of crowded curricula; possibly it would make little difference if students attended the odd course on good presentation and effective data display. But, at the very least, they could be encouraged to read a book such as this. Once sensitized to the pitfalls, they might be more prepared to put in the effort required to select a clear and effective means of presenting their results.

Consider even the most standard type of graph, with two axes, some tick marks

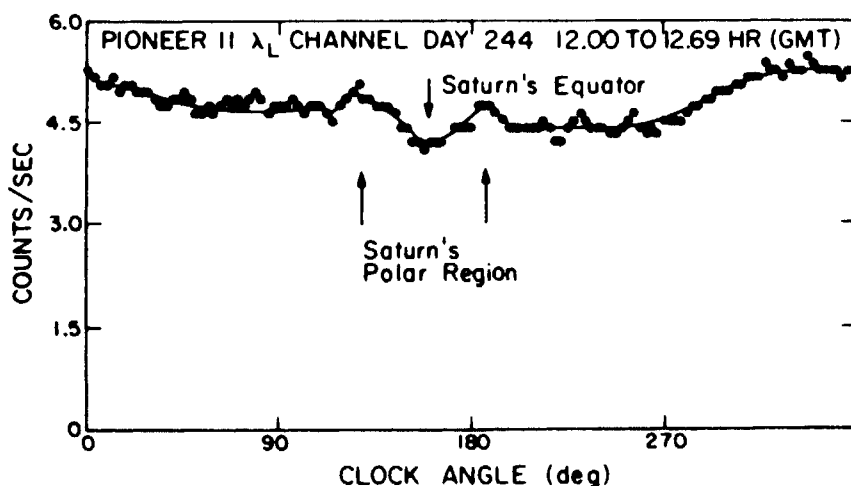
to show the scale, dots for the data, and lines connecting the dots. Do the data stand out, or are they obscured by the connecting lines? Do the data points overlap and obscure each other? Do the axes fall across data points? Are there too many tick marks? Do the labels clutter the picture and make it hard to grasp the significant features of the data? All these matters require proper attention. Other types of graphical medium, such as the scatterplot, pose other questions.

One especially fraught area is the display of multidimensional data. To graph the relations between ten or twelve different variables can stretch even the most active imagination. Most attempts fail in one way or another. The author suggests the use of colour to improve readability; this is fine in principle, but most journals refuse to include colour because of its cost. The scatterplot matrix — an array of coordinated plots of all possible pairs of variables, one against the other — manages to convey a great deal of information succinctly and clearly.

The danger with principles is that they can be taken too literally. A few horrific examples show what happens when the Darrell Huff Dictum, "always include zero in the scale", is overdone. To cramp the data into a tiny corner of an otherwise blank graph helps nobody. Huff was thinking more about advertising or political propaganda, and when preparing a graph for scientific publication, it should be permissible to assume that the reader has the intelligence to read the numbers on the axes.

If you're not willing to present your data effectively, then presumably you don't really care what they are. If you don't write papers that others can read, why write papers at all? The author's basic message is a simple one: remember that the important thing is the data. Granted this, all else will follow. □

Ian Stewart is a Reader in Mathematics at the University of Warwick, Coventry, CV4 7AL, UK.



The resolution of this graph is degraded by the inclusion of zero on the vertical scale.

A look around the kitchens

Alastair Hay

Controlling Chemicals: The Politics of Regulation in Europe and the United States. By Ronald Brickman, Sheila Jasanoﬀ and Thomas Ilgen. Cornell University Press: 1985. Pp. 344. \$38.45.

EARLY attempts by some United States Government Agencies to control the availability and use of chemicals were derided by critics, and the chemical industry in particular, as "cook-book" approaches to the subject. Fifteen years later things have changed, but analogy with cooking remains an appropriate way to describe the thesis of this book.

The authors have set out to compare how chemicals are regulated in Britain, West Germany, France and the United States. It is as if four cooks were asked to bake a loaf of bread — starting with the same ingredients, they follow different recipes but end up with a loaf nevertheless. The European cooks keep their techniques to themselves, arguing that their success lies in some secret mix. The American, on the other hand, has an audience to scrutinize all the ingredients, check the recipe and follow each step from mixing to baking.

The same applies to chemical regulation. Europeans prefer a consensus approach, negotiating with interested parties behind closed doors and presenting final versions of regulations which are rarely altered by Parliament. In the United States, however, the procedure is adversarial, with the government, Congress, the courts, public interest groups and industry all having a say in the matter.

While recognizing that the European pattern of involvement is more restrictive, the contention of Brickman and his colleagues is that, despite this lengthy and expensive process, legislation in Europe is as — or more — effective as that in the United States. (To prove the point, a comparative table of regulated substances is appended to the book, showing limits for the four countries.) Their conclusion is drawn from a thorough analysis of the European systems, some features of which, they argue, could well bear consideration by both regulators and legislators in the United States.

There are fundamental differences between the regulatory processes applied in Europe and the United States. Where the scientific evidence on the hazards — and the carcinogenic risk in particular — of a chemical is unclear, the Europeans tend to adopt a policy of 'innocent until proven guilty'. In the United States, however, guilt is assumed, until proven otherwise, and decisions about regulations are made from that standpoint.

Brickman *et al* point out that "concepts like 'reasonable' and 'practicable' sum up British regulatory philosophy, a philosophy generally shared by industry". Discussions about what these terms mean as far as particular chemicals are concerned are carried out in private between civil servants and industry. In the case of decisions taken by Britain's Health and Safety Executive, workers' trades unions are also involved. The result, say the authors, is usually a compromise that all sides are prepared to live with.

On the question of whether there is a 'safe' concentration for carcinogens to which workers can be exposed without risk, Britain seems to be out of step. British regulators, are favourably disposed to the concept of a threshold. As pointed out in the book, a guidance note issued by the British Health and Safety Executive in 1978, proposed a method of determining a "practical threshold of neoplastic response" while conceding that a "precisely defined" threshold cannot be attained. But Britain, it seems, is the only country that has banned some chemicals in the workplace because of their carcinogenic properties.

In Germany, the analysis of hazards is usually carried out by scientific advisory committees within each Ministry, who then also support appropriate regulations. It appears that these committees have a similar philosophy to the US Agencies in that they accept that there is no safe

threshold level for carcinogens. Because the German Commission on hazardous substances in the workplace argues that there is no way of assessing a threshold and calculating a tolerance level, this leaves the Federal Labour Ministry free to set exposure standards for carcinogens at any level it regards as economically and technically feasible. Similarly, the French are reluctant to establish threshold limiting values, and the strategy for regulating carcinogens is to "combine strict medical monitoring with reporting and personal hygiene requirements".

In the United States, however, legislation is preferred. But the strict standards for chemicals in the workplace recommended initially by Regulatory Agencies are invariably toned down by a process of attrition in the courts. The end result of this is standards similar to those in force in Europe.

While the author's principal conclusion is the need in the United States for a more uniform approach to rule-making, and an opportunity for more negotiations between competing interests, they do not envisage less information on chemical hazards being released to the public (something which would win widespread approval in Europe). It is ironic that at the same time that some Americans are recognizing the benefits of toning down their procedures, many Europeans are starting to see merit in a system of greater accountability — that is, they would like more people to be let into the kitchen. This well-researched and well-argued book will be an excellent starting point for people of both persuasions. □

Alastair Hay is a lecturer in the Department of Chemical Pathology, University of Leeds, Old Medical School, Thoresby Place, Leeds LS2 9JT, UK.

Get together

Marie Boas Hall

Science Reorganized: Scientific Societies in the Eighteenth Century. By James E. McClellan III. *Columbia University Press*: 1985. Pp.413. \$38.50, £41.80.

To join with others in a more or less formal group seems a natural human instinct of modern society. European scholars have long done so, in forms conditioned by the state of society: in the Renaissance under noble or princely patrons, later as a subsidiary of centralized monarchies, later still as a manifestation and a reflection of emergent nation-states. The first resultant academies were mainly literary or musical or scholarly; as science developed in the seventeenth century its practitioners drew together, either by correspondence or by association, and scientific academies, institutes and societies were born. And it is the societies and organizations themselves that form the focus of Dr McClellan's book.

Of these emergent academies the Royal Society of London (founded 1660, chartered 1662) and the Académie Royale des Sciences of Paris (1666) are the oldest in continuous existence, although not the first to be conceived. Both differed from earlier foundations in being independent of a private patron, in being incorporated entities and in publicizing the work of their members. The Académie (itself parallel to the Académie Française of 1635) was the prototype of modern national institutes and academies, its members paid by and sometimes appointed by the state, while the Royal Society was and is independent of the Government, though adviser to it and an agent of it.

Until roughly 1730 the Royal Society had the more eminent members and hence the greater reputation; slowly, in the wake

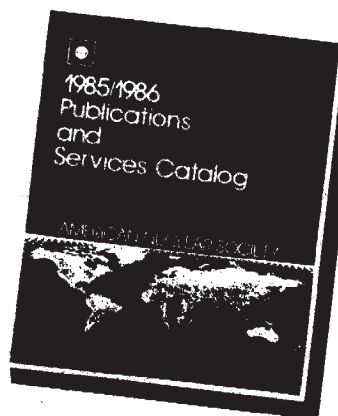
of reform in 1699, the Académie reversed the situation. Because of this and the fact that French culture and politics dominated eighteenth-century life in continental Europe, later academies were imitations of that of France, until Berlin, St Petersburg (both initially mainly consisting of foreign savants), Uppsala, Turin and a host of provincial cities all possessed formal organizations with charters, publications and paid members. These were sometimes purely scientific, sometimes, especially later in the century with growing linguistic nationalism, combining a scientific with a parallel literary section, while England, after the growth of private provincial societies, about 1790 saw the rise of the "Lit. and Phil." movement.

What was the cause of all this activity? Dr McClellan as he surveys the eighteenth-century scientific academy sees it as the result of increasing professionalism in science (he does not discuss non-scientific groupings). He is interested in the organizations, which he describes and analyzes fully, not in the individuals who made them up nor, except where cooperation between societies was involved, in the work they accomplished. This gives a lack of solidity to his account, reinforced by a certain naivety about European (including English) history, pointed up by the American sections. (Nor has he a very sharp grasp of language.) He is at his best in describing the structure of the new continental European academies and the relationships between them and he rightly regards as innovative the beginning of a formal (as distinct from a personal) exchange of publications begun in mid-century. There is a very great deal of information in this book which clearly shows how scientific academies developed during the eighteenth century. □

Marie Boas Hall, 14 Ball Lane, Tackley, Oxford OX5 3AG is Emeritus Reader in History of Science and Technology of the University of London.

PUZZLED? STOP GUESSING

FOR **FACTS** ABOUT
NUCLEAR SEND FOR YOUR
FREE PUBLICATIONS AND
SERVICES CATALOG



AMERICAN NUCLEAR SOCIETY
555 N. Kensington Ave.
La Grange Park, IL 60525

A double-blind test of astrology

from Shawn Carlson

Two double-blind tests were made of the thesis that astrological 'natal charts' can be used to describe accurately personality traits of test subjects.

ALTHOUGH there have been many published 'tests' of astrology, those with positive results (confirming the astrologers' thesis) have been largely dismissed by scientists on the grounds of technique. Those with negative results (disputing the astrologers' thesis) have been largely dismissed by astrologers on the grounds that they fail to test what the astrologers consider to be the essential aspects of their work. Indeed, astrologers complain that most scientific tests have tested the scientist's concept of astrology, not astrology as practised by the 'reputable' astrological community. Both criticisms may be valid.

My purpose has been to avoid these criticisms by designing an experiment that would meet the tight specifications of both the scientific and astrological communities. Such an experiment was designed with the help of scientists, statisticians and astrologers. We decided to test what we shall call (for simplicity) the 'fundamental thesis of natal astrology' as the proposition that:

the positions of the 'planets' (all planets, the Sun and Moon, plus other objects defined by astrologers) at the moment of birth can be used to determine the subject's general personality traits and tendencies in temperament and behaviour, and to indicate the major issues which the subject is likely to encounter.

The device used by astrologers to make predictions is called a 'horoscope', in essence a picture showing the positions of the various astrological objects in the heavens on a backdrop of twelve equally-spaced imaginary sectors called 'houses', as seen from a particular place and time on Earth. Typically, a horoscope includes a table which shows the angular relationships (or 'aspects') between the astrological objects. If the place and time are those of a person's birth, the horoscope is called a 'natal chart' (see Fig. 1), from which astrologers derive information about a subject's personality and character. The descriptive text thus derived is called a 'natal chart interpretation'.

To satisfy both the scientific and astrological communities, we chose as advisers people held in high esteem by their respective communities. The astrologers helped us to formulate the proposition given above as central to 'natal astrology' (the subfield of astrology which deals with birth data) and yet scientifically testable. Care was taken to include all suggestions by the astrologers provided they could be followed without biasing the experiment

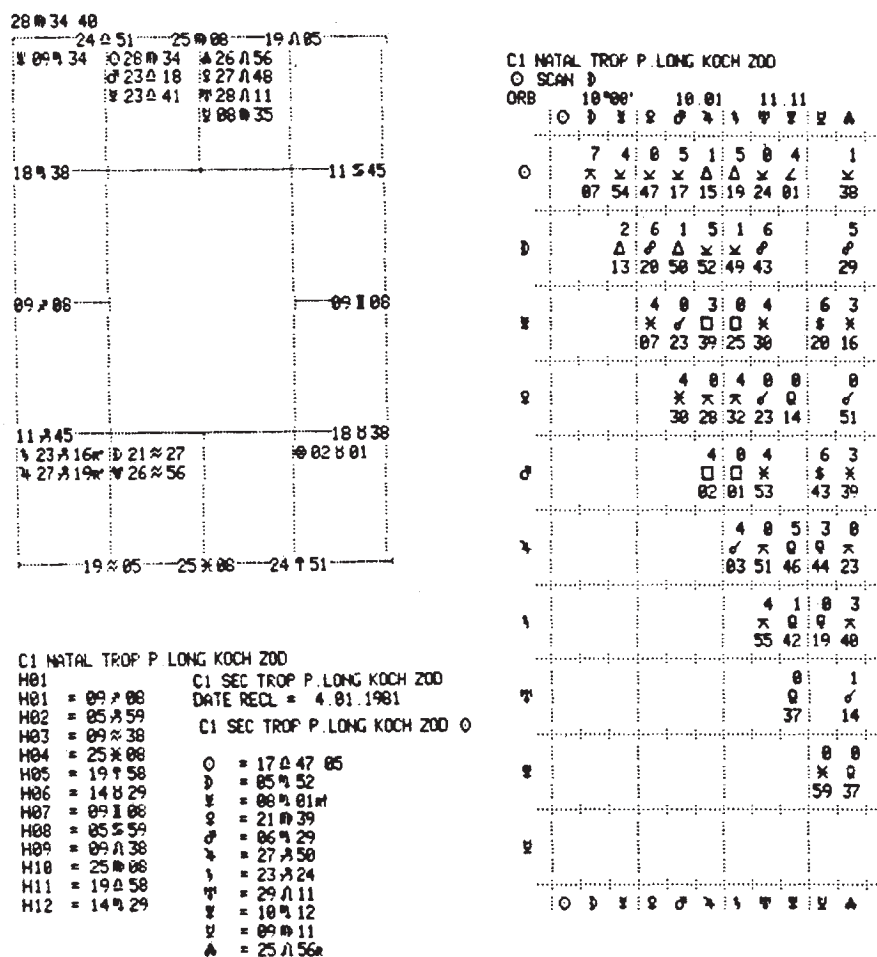


Fig. 1 Computer derived 'natal chart', showing positions of astrological objects as seen from the time and place of a person's birth.

for or against the astrological thesis. We took great care to eliminate all biases which could tend to 'randomize' the results and thus favour the scientific hypothesis over the astrological one. Similar care was taken to make sure that hidden clues were not available which could be used by astrologers or subjects to choose correct answers not based on astrological information alone.

The experiment designed by these means consists of two parts.

Part 1. Volunteers provided information from which their natal charts and interpretations were constructed by astrologers. Each subject then attempted to select his own natal chart interpretation from a group consisting of his own and two other interpretations chosen at random from the

whole group. The subjects made first and second place choices; ties were not allowed. Subjects were also asked to rate each interpretation on a 1-10 scale (10 being highest) as to how well each interpretation fit them. If their selections are random (scientific hypothesis), we would expect them to select their own interpretation one third of the time. The astrologers predicted, given the design of the experiment, that the subjects would be able to choose their own interpretation "at least half" of the time.

Part 2. The participating astrologers were separately given the natal chart of a random subject and an objective and respected measure of his personality traits called the California Personality Inventory (CPI). They were also given two other

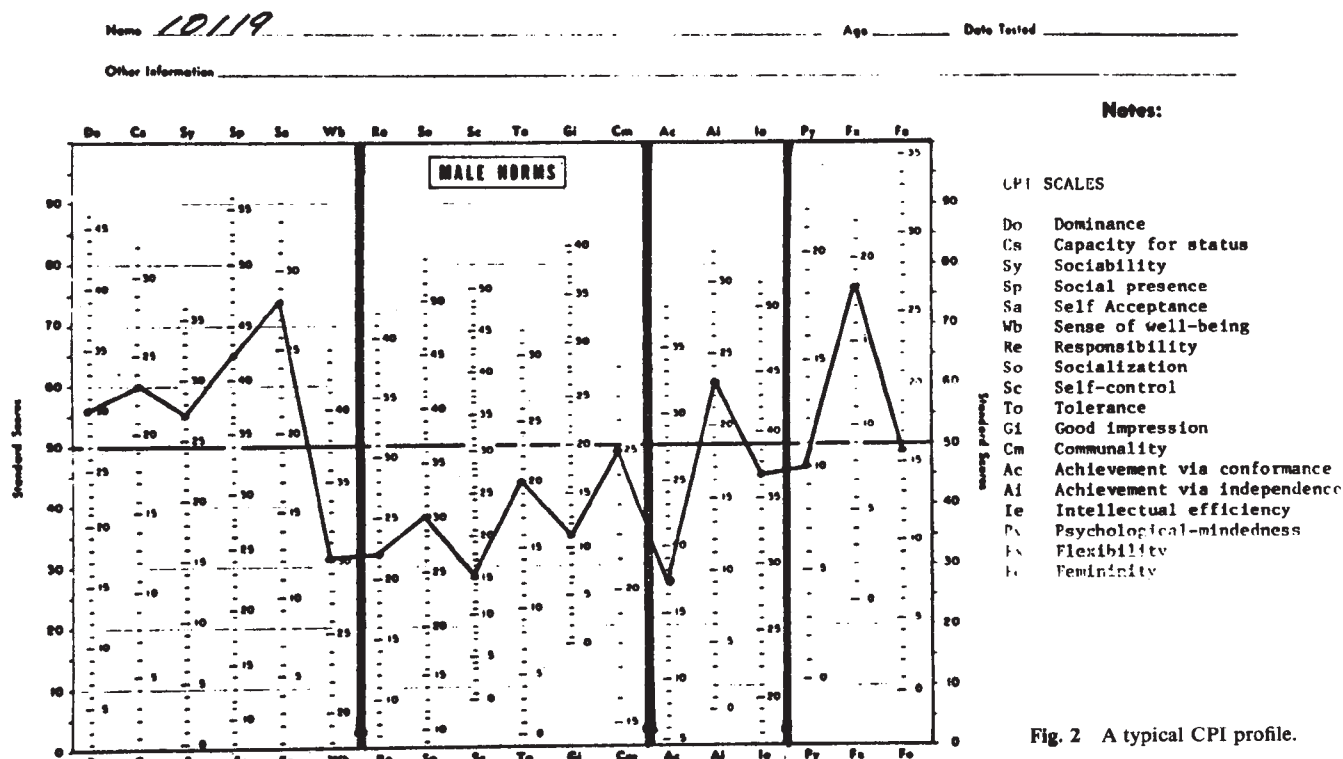


Fig. 2 A typical CPI profile.

CPIs chosen at random from the group of all the subjects' CPI test results. The astrologers were then asked to select the two CPIs (first and second choice, no ties allowed) which described personalities closest to the personality indicated by the natal chart. They also rated each CPI on a 1-10 scale (10 being highest) as to how closely its description of the subject's personality matched the personality description derived from the natal chart. The scientific hypothesis predicts a correct choice one third of the time; the astrologers predicted a correct choice half the time or more.

The two parts of the experiment are complementary; the first is not sensitive to biases in the CPI, the second does not make the possibly false assumption that subjects can accurately judge their own personalities. We decided at the outset to require a 2.5 standard-deviation increase over random chance to interpret the results as favouring the astrological hypothesis. Similarly, a disagreement by 2.5 standard-deviations or more would be required to reject the astrological hypothesis in favour of the scientific one. Otherwise, we decided in advance, we would draw no conclusions of significance. No data were analysed until all the data had been collected.

Experiment design

To eliminate bias, both anticipated and unknown, we made extensive use of double blind techniques. All subjects were assigned a five-digit random code number. Neither the astrologers nor the experimenter knew what code number corresponded

to which person. These lists were solely under the supervision of Richard A. Muller, Professor of Physics at the University of California, Berkeley.

Guidance was sought both from the scientific and astrological communities. To help ensure correctness of the testing method and statistical analysis, the scientific adviser was Professor Muller. So that participating astrologers should be respected by the astrological community, we sought the advice of the National Council for Geocosmic Research (NCGR), an organization which has been involved in much astrological research in the past and which has the respect of astrologers world wide; NCGR nominated persons who consented to be our astrological advisers, and who carefully reviewed the experimental design and made many suggestions.

After they were satisfied that the experiment was a 'fair test' of astrology, our astrological advisers established their predictions (50 per cent for both part one and part two) as the minimum effect they would expect to see. They also compiled a list of approximately 90 astrologers with some background in psychology who were familiar with the CPI and held in high esteem by their peers. It was the opinion of the advisory astrologers that a random sample from this list would be able to score at the predicted 50 per cent level. All were invited to participate; 28 accepted. (Only two astrologers who participated were not on the original list. They heard of the experiment and wanted to take part. After their qualifications had been vouched for by NCGR, they were admitted.)

Constructing a natal chart is a simple but laborious mathematical process, so computers are well suited for the task; several machines designed specifically for this purpose are available on the market. To save time and ensure accuracy, all natal charts were constructed by Mr. Caveney (President of the San Francisco chapter of NCGR) and Mr. Nelson (Secretary of the San Francisco chapter) on a Digicom DR 70 Astrological Computer; they were spot-checked by hand calculation.

The California Personality Inventory (CPI) is a standard personality test¹⁻³ which has been used extensively since 1958. It was chosen over other available personality tests because the advising astrologers judged the CPI attributes to be closest to those discernable by astrology. By choosing this test we were thus trying to maximize the ability of the astrologers to match CPI data with natal charts without introducing a pro-astrology bias. Other experiments have been done using the CPI with apparently positive results⁴.

The CPI consists of 480 true-false questions, each of which helps to rank a subject on one of 18 personality attribute scales (for example, dominance, passivity, femininity, masculinity). The subject's score on each scale is compared to the norm for that scale. The scores can be plotted on a graph (see Fig. 2) which readily conveys this information. Such a graph is called a 'CPI profile'.

Personality tests were graded, after names had been replaced by code numbers, by volunteers (undergraduate students) who were in no other way connected with the experiment. From spot

checks of the grading, we determined (95 per cent confidence level) that mistakes by the graders contribute an error of more than two points to CPI scores on fewer than 2.6 per cent of the individual scores, an insignificant effect.

Subjects were solicited by advertisements in San Francisco Bay area newspapers, classroom announcements and postings on and off the Berkeley campus. (To protect the confidentiality of the data and the rights of subjects, all procedures were checked by the University of California Office of Fair Treatment to Human Subjects before beginning data collection.) Approximately 70 per cent of the subjects were college students and about one half of these were graduates. All subjects were required to fill out a questionnaire with their natal data (birthday, including exact time and location of birth). They were also asked whether they (1) believe in astrology, (2) believe somewhat, (3) have no opinion, (4) disbelieve or (5) strongly disbelieve in astrology, and to state whether they had ever had a natal chart constructed before. Subjects were not told that these questions affected subject selection, but those who chose "(5) strongly disbelieve" were eliminated, on the grounds that this opinion might bias them, either consciously or unconsciously, against selecting the interpretation which best fitted them. All those who had previously had a chart constructed were similarly eliminated because they might be able to select (or reject) the correct interpretation based on a knowledge of what to expect. Strong believers who had never had their charts done were, however, not eliminated; this belief alone could not help them select the correct interpretation. All subjects had to be at least 17 years old. Failure to take the CPI resulted in rejection.

To avoid the possibility that a subject may have had his natal chart constructed elsewhere, or may have changed his opinion about astrology, between the time he submitted his natal data and the time he was given the final natal interpretations (typically 8–10 weeks), all subjects were required to fill out a new questionnaire before being asked to choose their own natal interpretations in Part 1 of the experiment. Two subjects were eliminated at this point, one who admitted to being a professional astrologer (and who had apparently lied on the first questionnaire) and another whose opinion of astrology had changed from "disbelieve in astrology" to "strongly disbelieve in astrology".

We encouraged prospective subjects to participate by promising them a copy of their natal chart, CPI test results and interpretation, the completed natal interpretation, and a copy of the final results of the experiment.

Although we required a departure from random of only 2.5 standard deviations to interpret the results as favouring the astro-

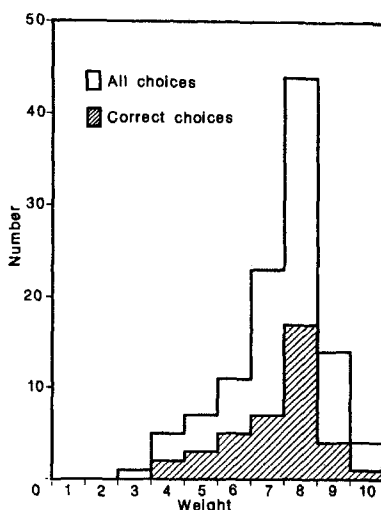


Fig. 3 Histogram showing weights assigned by astrologers to the CPI profiles they felt best fit the natal charts. CPI profiles rated higher are not more likely to be correct.

logical thesis, we originally planned to be able to distinguish between the two hypotheses at the four standard deviation level. This is why the number of subjects required was chosen to be 128, with a further 128 as a control group. Thus the total number of subjects was originally 256, but many of these did not complete all phases of the experiment.

Many lost interest and did not return their data to us. Some moved in the time taken to send them the test materials and did not leave a forwarding address. Two room-mates became emphatically convinced that astrology was the work of the devil, and refused to continue in what they called 'an experimental test of evil'. We were forced also to eliminate 12 subjects either because they did not follow directions correctly or did not return all the needed materials to us.

The use of double blind techniques is most important during this stage of the experiment. During the process of rejection of data, the experimenter had no access to any information that might introduce bias. In the end, only 177 subjects (83 test group, 94 control) remained for Part 1 of the experiment. Neither were we able to collect all the data we had hoped to in Part 2 of the experiment. First, fewer astrologers than hoped for agreed to participate. 224 data envelopes were mailed to only 28 astrologers, some of whom simply refused to participate as promised. Some declined after they discovered how much time was required on their part. One tried to bargain his services in exchange for free access to our raw data, and declined to participate when his terms were refused. For these reasons, we obtained only 116 usable subjects for Part 2 of the experiment. The large reduction in numbers was unanticipated and reduced the expected discrimination between hypotheses for Part 1 to 3.2 stan-

dard deviations, and for Part 2 to 3.9 standard deviations. However we do not believe that the loss of data could bias the results of the experiment in any significant way.

Bias and control

Experiments using human subjects are subject to a special class of biases which do not normally have to be considered by a physical scientist. An experiment must be designed so that the psychology of the subjects will not alter the results. The major potential biases which required specific control in the experiment design were as follows:

Sun-sign bias. By the astrological definition, the 'Sun-sign' refers to the constellation of the zodiac in which the Sun resides when the person is born. If the Sun-sign should play an important role in the average chart and if people are generally familiar with the characteristics of their Sun-sign (through newspaper horoscopes, for example), we might expect them to select the correct interpretation at a better-than-chance level regardless of whether or not the astrological hypothesis is correct.

To correct for this, each member of the test group was matched to a member of the control group born under the same Sun-sign. Following the astrologers' recommendations, we required that the age difference between these subjects be at least three years, so that their natal charts would be 'sufficiently dissimilar', but otherwise, the assignment was made randomly. Both test and control subjects were given the same three interpretations.

If the astrological hypothesis is false, members of both groups should identify the test subject's interpretation with equal frequency. If the hypothesis is true, the test group should score significantly higher than the control.

We believe that many other sources of bias are eliminated by the design of the experiments and by the standard formats in which data was conveyed to and from the astrologers. The possibility that subjects (in Part 1) would be tempted to choose flattering interpretations (or the opposite) should not bias our conclusions, given that we are comparing hits and misses between test and control groups. We sought to eliminate from the interpretations clues to their origin that might guide a subject to a correct choice, such as astrological terms that might be part of subjects' general knowledge, and to eliminate from the data with which astrologers worked, information that might allow them, consciously or otherwise, to bury hidden clues in their interpretations. Thus we eliminated from the charts transmitted to the astrologers information about the time and place at which subjects had been born. (Although the charts in principle allow the reconstruction of this information, the process is time-consuming, and the outcome unlikely to be of direct help to

astrologers.) We did not tell the astrologers of the gender of the subjects, partly because gender-specific elements of the interpretations might lead to bias, but also because the CPI scores cover different ranges for males and females.

To eliminate the possibility that subjects could pick up clues other than the astrological information we were testing, and to insure that the information given to subjects was as uniform, and thus as comparable as possible, the interpretations followed a predetermined format designed to specify what factors astrologers should derive from the chart and to set a limit on the length of written material. The format was developed in collaboration with the advising astrologers. The specific categories which astrologers were required to address were: (1) Personality/temperament; (2) relationships; (3) education; (4) career/goals; and (5) current situation. The astrologers typed each interpretation on pages supplied by the experimenter and containing the proper headings, again to keep the interpretations as uniform as possible.

The format also specifies: (1) that advice or predictions were not to be given, on the grounds that such information could not help the subject to select the correct interpretation but that it might well lead him to discard an accurate description because he disagrees with the advice or predictions given. (2) That no direct reference to the chart was to be made (e.g. "you have sun in Leo"). (3) That no information relating to the subjects' ages was to be given.

Subjects were asked to rate each section of each natal interpretation on a 1-10 scale. Then they were asked to write down, for each section, the code number of the interpretation which fitted them best and second best.

One complication of our study arises because a subject's ability to select the correct description of himself from a given group must depend on how well he knows himself. If people generally have an inaccurate self-image, one would not expect subjects to select the correct interpretation no matter how accurate are the results of astrology. We devised the following scheme to understand this potential bias.

The CPI is generally accepted by psychologists as a moderately accurate description of a person's personality. Each test subject was given his own CPI profile and two others randomly selected from the group. He was then asked to select the profile which he felt best fitted him. Each subject was provided with the following: (1) three sample CPI profiles; (2) a synopsis of what high and low scores in each category tend to be for males and females; (3) a letter explaining about a CPI profile and how to go about making the selection.

To control for possible psychological bias, we elected to use the same test and control groups as in Part 1, but since the CPI is graded on different scales for males and females, we had to match male

Scientists wishing to familiarize themselves with astrology as astrologers practise it may find the following works useful:

- Gleadow, R. *The Origin of the Zodiac* (Castle, Pasadena, 1968).
Dean, G. *Recent Advances in Natal Astrology* (Analogic, 1977).
Rudhyar, D. *The Astrology of Personality* (Doubleday, New York, 1970).
Meyer, M. *A Handbook for the Humanistic Astrologer* (Anchor, Manhattan Beach, California, 1974).
Keyes, K. *Master Guide to Preparing Your Natal Horoscope* (Parker, Los Angeles, 1984).
Jones, M. *The Sabian Symbols in Astrology* (Shambhala, Berkeley, 1953).
George, L. *A to Z Horoscope Maker and Delineator* (Llewellyn, St Paul, Minnesota, 1954).
Gauquelin, M. *The Cosmic Clocks* (Regnery, Chicago, 1967).
Gauquelin, M. *The Scientific Basis of Astrology* (Stein and Day, New York, 1966).
Dean, M. *The Astrology Game* (Beaufort, Beaufort, California, 1980).

(female) test group members to male (female) control members. Thus, the test-control group assignments had to be re-established.

Obviously, since the natal chart depends entirely on the natal data, inaccuracies in the latter would produce inaccuracies in the former. The astrologers insisted that the birth time be accurate to within 15 minutes. In order to assure this, when the subjects took the CPI, they were obliged to show documentation of their natal data including, especially, birth time. Although we preferred birth certificates, hospital and county record or other 'official' documentation, we also accepted baby books provided that the birth time was recorded when the child was born. (A more complete discussion of biases may be found in L.B.L. Preprint No. 20480.)

Double-blind procedures

An important difference between this and many previous tests is our extensive use of double-blind techniques. The important procedures we followed are outlined below.

As soon as subjects had taken the CPI, their questionnaires were put into alphabetical order and given to an assistant, who assigned a five random-digit code number to each in turn. No two subjects were assigned the same code number. The assistant then filled out three 3 × 5 index cards for each subject. On the first he put

Name-Code Number; these he filed in alphabetical order. On the second he put Code Number-Name; these were filed in numerical order. The purpose for these cards was for easy record-keeping. The cards were maintained under the supervision of Prof. Muller, and could be released to the experimenter only with his consent. At no time during the data collection did the experimenter have access to any information relating subjects' identities to code numbers. This control was abandoned only when all the data had been collected and the methods of analysis had been established.

The assistant also made a third set of cards each containing the code number of a particular subject and his natal data. These cards were given to astrologers Michael Caveney and Chris Nelson in envelopes unopened by the experimenter, for the construction of natal charts. Since the CPI was given at three different sessions (typically about 4 weeks apart), not all the assignments were made at the same time, nor did the astrologers receive all the natal data cards at one time. They did receive, however, the natal data cards within five days of the CPI testing dates.

Subjects who failed to show documentation of birth time, date and location when they took the CPI were automatically put into the control group (there were 43 such subjects). All questionnaires in each group were sorted into twelve Sun-sign groups. We then randomly assigned subjects to the control group until the number of control group members equalled the number of subjects remaining in each Sun-sign group. The remaining subjects comprised the test group. If there was an odd number of subjects in a Sun-sign group, the odd person was put in the control group.

The questionnaires in each Sun-sign group were then shuffled thoroughly and the names and birthdates were listed in the order in which they landed. As stated earlier, we required the birth dates to be at least three years apart to ensure that the natal charts were sufficiently dissimilar. The first test subject of each Sun-sign group was then matched to the first control subject of the same group whose birth date was far enough apart. The procedure was repeated for the second, third, etc. test subjects until we were left with test and control subjects whose birthdates were within three years of each other (which happened in all 12 Sun-sign groups).

To include these remaining subjects, we were forced to do some rematching in the

Table 1 Data from subject selections of natal chart interpretations

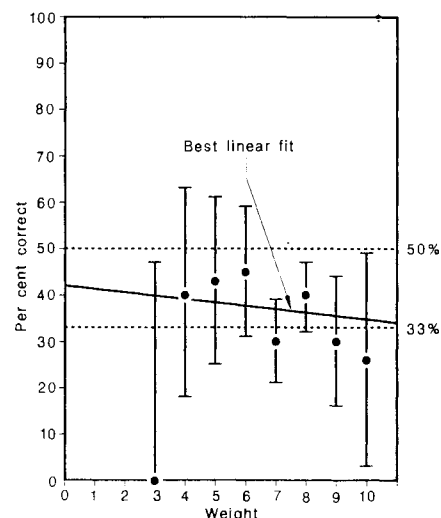
	Total	First choice	Second choice	Third choice	
Test group	56	25	16	15	18.67 ± 3.53
Control group	50	21	13	16	16.67 ± 3.33
Test group	83	28	33	22	27.67 ± 4.29
Control group	94	42	34	18	31.33 ± 4.57
					CPI PROFILE SELECTION INTERPRETATION SELECTION

following way. We started at the top of the control group and went down until we found a birthdate which satisfied the requirements that it was at least three years away from the unmatched test person and that the birthdate test group member to which it was attached was at least 3 years away from one of the unmatched controls. We then rematched the originally-matched control group member with the previously unmatched test subject, and matched the unmatched control group member with the previously matched test member. This was continued until all the test and control group subjects were matched. Because seven Sun-sign groups had an odd number of subjects, there remained seven unmatched control group members.

For the second part of the experiment, male (female) test members had to be assigned to male (female) control members. Thus, the above test subject to control subject matching had to be redone. Since the test and control groups were randomly chosen, no changes in them were made. However, all the matchings changed as male (female) test subjects were randomly matched to male (female) control subjects.

During the conduct of the experiment, each subject was given two envelopes, one containing the materials needed for the selection of the natal chart interpretation and the other containing the materials for the selection of the CPI profile. The natal interpretation envelope contained: (1) three natal interpretations; (2) a pretyped sheet on which the subjects were to detail their choices; (3) a questionnaire asking their opinion of astrology and whether they had had their chart done before (See Experiment Design); and (4) a letter explaining how they were to go about making the selections. The CPI envelope contained: (1) three CPI profiles; (2) a preformatted sheet on which the subjects were to detail their first and second place choices, plus 1-10 ratings; (3) a summary of what the CPI scores mean; and (4) a letter explaining how they were to go about making their selections. Since we had labelled the natal charts with the code number of the person for whom they were constructed, we could not do the same with the CPIs, since the subjects might recognize a code number appearing twice. The CPIs were first labelled with the code numbers of the persons to which they corresponded, and then relabelled by finding

Fig. 4 Graph showing percentage correct versus rating for astrologers first place choices in CPI profile natal chart matching. The best linear fit is consistent with the scientifically predicted line of zero slope. No significant tendency for the astrologers to be more correct when they rate a CPI as highly matching a natal chart.



$\ln(1/x)$ for each code number, x , on a hand calculator and taking the last 5 digits on the calculator display as the new code number. All natal interpretations and CPI profiles were put in numerical order in the envelopes.

The astrologers were sent materials in two separate mailings. In the first, each received: (1) the number of natal charts he had agreed to interpret when he chose to take part (typically 4); (2) a copy of the format by which the charts were to be interpreted; (3) the paper with headings on which the interpretations were to be typed; (4) a letter explaining the symbols used on the computer-constructed natal charts, deadlines and notice of when they might expect the materials for part two of the experiment; (5) a postage-paid return envelope. After they had returned the natal charts and interpretations, the astrologers received the second mailing containing: (1) the number of natal charts (plus three CPIs for each natal chart) which they had agreed to match to CPI profiles; (2) a copy of *The Interpreter's Syllabus for the CPI*² (a booklet explaining all the CPI attributes and how to interpret them in detail); (3) a preformatted sheet on which the astrologers were to detail their first and second choices and ratings; (4) a letter explaining how to go about making the choices; (5) a postage-paid return envelope for the data.

To save the astrologers work, they were allowed to make the CPI matchings to the natal charts they had already interpreted. They were also typically sent an additional natal chart and CPIs to match. To ensure that the astrologers would not mix up the CPI profiles between natal charts, the three profiles for each natal chart were ordered randomly, then labelled *a*, *b* and *c* respectively.

A total of 226 natal charts were sent out to be matched with CPI profiles. Of these, only 15 had no documentation of birth time. None of these 15 were returned by the astrologers to be included in our data.

Results

Part 1. Subject selection of natal chart interpretations. After a predetermined data collection period of 10 weeks, we had astrological data from 83 test and 94 control subjects and CPI self-selection data from 56 test and 50 control subjects. The data are displayed in Table 1.

The test group selected the correct interpretation as its first choice at the rate of 0.337 ± 0.052 , the control group at the 0.447 ± 0.049 rate, 2.34 standard deviations above chance. Although this fluctuation is less than 2.5 standard deviations, the level we had chosen to call 'significant', it does require comment. Since this fluctuation occurred in the control group and control subjects were not given their own interpretations, this cannot be interpreted as a possible astrological effect. Neither can it be correctly attributed to Sun/sign bias, since the test group did not score near the same level. We thus interpret this as a statistical fluctuation. The test group chose the correct interpretation as second best, describing them at the 0.398 ± 0.052 rate while the control group did so at the 0.362 ± 0.049 rate. Finally, the correct interpretation fell as the test group subjects third choice at the 0.265 ± 0.052 level and, for the control group, at the $0.191 \pm$

Table 2 Data from astrologers matching natal charts to CPI profiles

	Total (n)	Chance (n/3) [Expected s.d.]	Astrologers predicted (n/2) [Expected s.d.]	No. of correct CPI chosen	Standard deviation away from 0.35	Standard deviation away from 0.50
First choice	116	38.5 [5.1]	58.5 5.4	40	+2.56	3.34
Second choice	114	38.0 [5.0]	None	46	+1.48	—
Third choice	114	38.0 [5.0]	None	28	2.0	—

The data are consistent with chance, inconsistent with astrological hypothesis.

0.049 rate. All this is consistent with the scientific hypothesis.

When the first few data envelopes were opened, we noticed that on any interpretation selected as a subject's first choice, nearly all the subsections were also rated as first choice. We then realized that we had no way of guaranteeing that subjects were rating each section of the interpretations independently of others they had already read. Without such a guarantee, spurious results favouring either hypothesis could have easily appeared. So we rejected these data as not having been collected under the proper controls.

Next we looked to see how well the subjects were able to select the correct CPIs. The test group selected the correct CPI as their first choice at the 0.446 ± 0.063 rate, while the control did so at the 0.420 ± 0.066 rate, showing no significant difference between the two groups. The test group chose the correct CPI as the second place choice at the 0.286 ± 0.063 rate, while the control did so at the 0.260 ± 0.066 rate, again no significant difference between the two groups. Finally, the test group chose the right interpretation as their third choice at the 0.268 ± 0.063 rate while the control group did so at the 0.320 ± 0.066 rate. Once more, there is no significant difference between the two groups.

Part 2. Natal Chart-CPI matching. A total of 116 data envelopes were returned by the astrologers. In the 116 envelopes there was a total of 116 first place choices, 114 second place choices and 320 ranked choices (weight factors indicating how well the astrologers felt the natal chart matched each CPI, on a scale of 1 to 10) (Table 2).

The data were first analysed without taking the 1-10 weight factors into account. The astrologers selected the correct natal chart as their first place choice at the 0.34 ± 0.044 rate, in agreement with the scientific hypothesis of 0.33 and in disagreement with the astrological hypothesis of 0.5 by 3.3 standard deviations. The correct CPI was chosen as the second place choice at the 0.40 ± 0.044 rate which is also consistent with the scientific hypothesis. (The astrologers had made no firm prediction about the second place choice.) The correct CPI was chosen as the third place choice at the 0.25 ± 0.044 rate, again consistent with the scientific hypothesis.

Next we took the weights into account, by a method established before studying the data. (The establishment of methods before data analysis is important in order to prevent the subtle bias that comes from selection of analysis procedures.) We first made a histogram of the weights which the astrologers assigned to all their place choices, regardless of whether or not those choices were correct (Fig. 3). The data are sharply peaked at a weight of about 8. The second histogram in Fig. 3 shows the ratings of only those first place choices which

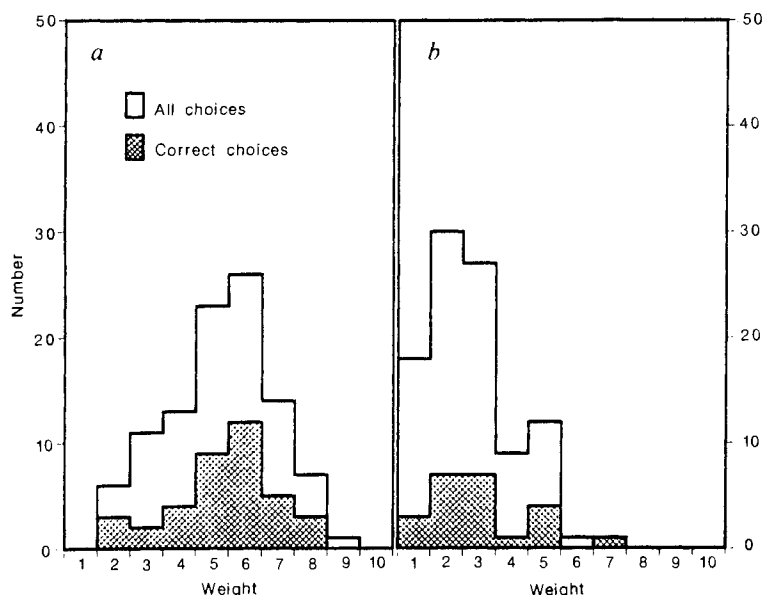


Fig. 5 Histograms showing rates assigned to astrologers' second (a) and third (b) place choices.

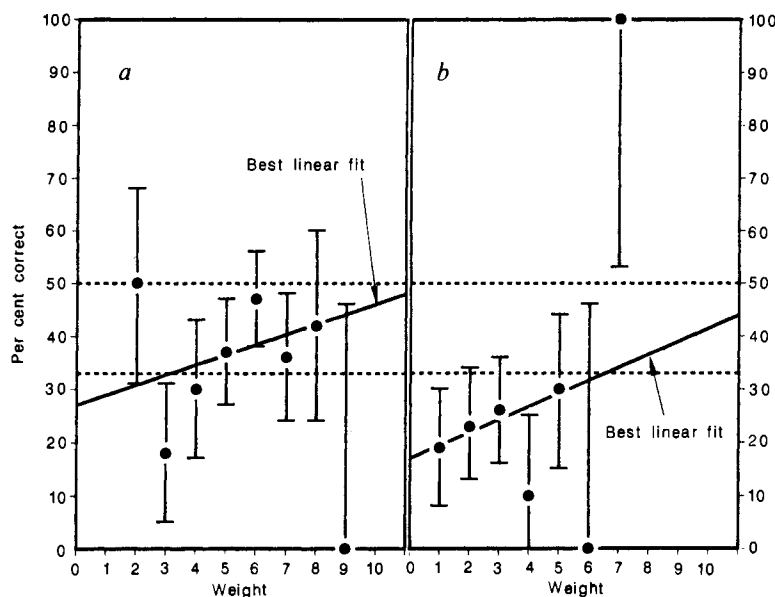


Fig. 6 Percentage of correct CPI profiles versus rates, chosen by astrologers as their second (a) and third (b) place choice. Best linear fits are consistent with chance.

were correct. If the astrological hypothesis were true, one might expect the correct first place choices to have higher weights on average than the whole group of first place choices. Thus, the new histogram should be skewed to the right. On comparing the two histograms, however, we see that they are very similar; no such skewing appears.

The scientific hypothesis predicts that 1/3 of the choices at any weight should be correct choices. Figure 4 shows the percentage correct for each weight with the appropriate error bars, and the best linear fit with slope -0.01 ± 0.02 . The slope is consistent with the scientific prediction of zero slope. The same analysis on the second and third place choices yields Figs 5a, b and 6a, b. The slope of the best linear

fit to the data on Fig. 6a is 0.019 ± 0.02 while that of Fig. 6b is 0.0026 ± 0.02 both consistent with the scientific hypothesis (zero slope).

Conclusions

From the results of Part 1 (subjects selecting interpretations), we notice that the test group scored at a level consistent with chance and within 2.5 standard deviations of the control group. The large (2.34 s.d.) but not significant (less than 2.5 s.d.) fluctuation in the control group is attributable to statistical fluctuation, not to a Sun-sign bias. These results are consistent with the scientific hypothesis. However we cannot use the result to rule against the astrological hypothesis, because the test subjects were also unable to select their

own CPI profile at a better-than-chance level. At the 95 per cent confidence level, the test subjects were unable to select their own CPI profile at better than the 0.57 rate. There are many reasons which could explain why the test subjects were unable to select the correct CPIs at a higher rate:

- (1) Subjects may have had difficulty relating to the graphical presentation of the CPI information.
- (2) Some subjects may have recognized correct information about themselves, but subconsciously chose a CPI which did not describe them as well to avoid admitting they have certain character traits. Such denial in a large percentage of the subjects would tend to cancel a positive effect.
- (3) The CPI may not test the kind of attributes by which subjects may easily recognize themselves.
- (4) People may be unable to recognize accurate descriptions of themselves.

Our experiment does not distinguish between these possibilities. Professor H. Gough (author of the test and respected experimental psychologist) is familiar with nearly all published experiments using the CPI. At our request he searched through the literature for any experiment demonstrating the ability of test subjects to recognize accurate descriptions of themselves. To his and our knowledge, no other test of this kind has ever been done. Thus, we believe there exists presently no scientific evidence from which one can conclude that subjects can select accurate descriptions of themselves at a significant rate.

If subjects cannot recognize accurate descriptions of themselves at a significant level then the experiment would show a null result no matter how well astrology worked. On the other hand, any astrological effect demonstrated in this way would require a consistency check. One would have to see if subjects could recognize the kind of information astrologers give them about themselves, which was derived in a manner *known* to be reliable. Thus, until and unless such a self-recognition ability can be shown, we conclude that subject selection of astrologically derived information is a poor test of astrology. (This is a problem in approximately 30 per cent

of all experiments which claim a significant astrological effect.)

The conclusions to be reached from Part 2 (CPI-natal chart matching) of the experiment are somewhat more illuminating. What is striking about these data is how poorly the astrologers performed, when their performance is compared to their predicted rate. It is consistent with chance, and is at the very significant 3.3 s.d. level below the astrologers' prediction. This is well beyond the 2.5 s.d. requirement we established before the beginning of the experiment as sufficient to refute the astrological hypothesis.

Before the data had been analysed, we had decided to test to see if the astrologers could select the correct CPI profile as either their first or second choice at a higher than expected rate. The scientific hypothesis predicts the CPI will fall in the first or second choice 66 per cent of the time. The astrologers did not make a specific prediction as to what they expected the rate to be. If the correct CPIs are chosen in the first and second place choices, then they will be depleted from the third place choice. Since the rate at which the astrologers chose the correct CPI as their third place choice was consistent with chance, we conclude that the astrologers were unable to choose the correct CPI as their first or second choices at a significant level.

In Fig. 4 the data are clearly inconsistent with the 'at least' 0.5 level predicted by the astrologers. Nor do the data suggest that the astrologers are more likely to be correct when they rate a CPI as well fitting the particular natal chart than they are when they weight it as poorly fitting the natal chart. The data appear randomly scattered about the 0.33 line and is hence consistent with chance. The scientific hypothesis predicts a line of zero slope, consistent with the slope observed. Figures 5 and 6 likewise show no convincing evidence that the astrologers tended to rate the correct CPIs higher than the incorrect CPIs.

We are now in a position to argue a surprisingly strong case against natal astrology as practised by reputable astrologers. Great pains were taken to insure that the experiment was unbiased and to make sure that astrology was given

every reasonable chance to succeed. It failed. Despite the fact that we worked with some of the best astrologers in the country, recommended by the advising astrologers for their expertise in astrology and in their ability to use the CPI, despite the fact that every reasonable suggestion made by the advising astrologers was worked into the experiment, despite the fact that the astrologers approved the design and predicted 50 percent as the 'minimum' effect they would expect to see, astrology failed to perform at a level better than chance. Tested using double-blind methods, the astrologers' predictions proved to be wrong. Their predicted connection between the positions of the planets and other astronomical objects at the time of birth and the personalities of test subjects did not exist. The experiment clearly refutes the astrological hypothesis.

I thank Professor Richard Muller for providing funding from his Alan T. Waterman Award, and for invaluable assistance and guidance; Professor Doksum, Statistics, UC Berkeley, for acting as an advisor at the conception of the experiment; Tony Joseph for many helpful conversations and for being our astrological advisor; Michael Caveney and Chris Nelson for constructing the natal charts and being astrological advisors; the participating astrologers for their expertise, time, and effort; Terry Mast, Teresa Weed, Geoffrey Dean for helpful criticisms; Kent McArthur who first suggested the use of personality tests; Susan Miller, who assisted in the advertising, and to those Berkeley undergraduates who gave up their time to act as assistants and CPI test graders.

Shawn Carlson is at the Department of Physics, University of California, Berkeley, California 94720, USA. Present address: Department of Physics, University of California, Los Angeles. Address for correspondence: Bldg 50, Rm 348, Lawrence Berkeley Labs, Berkeley, California 94720, USA.

Received 11 April 1983; accepted 14 October 1985.

1. Gough, H. G. *Manual for CPI* (Consulting Psychologists Press: Palo Alto, 1957).
2. Gough, H. G. *Advances in Psychological Assessment* Vol. 1, (ed. McReynolds, P.) (Science and Behavior Books, Palo Alto, 1968).
3. Megargie, E. I. *The CPI Handbook* (Jossey-Bass, San Francisco, 1972).
4. Pellegrini, R. J. *J. Psychol.* **85**, 21-28 (1973).

Direct observation of He^+ pick-up ions of interstellar origin in the solar wind

E. Möbius*, D. Hovestadt*, B. Klecker*, M. Scholer*, G. Gloeckler† & F. M. Ipavich†

* Max-Planck-Institut für Physik und Astrophysik, Institut für extraterrestrische Physik, 8046 Garching, FRG

† Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742, USA

Singly-ionized helium with a velocity distribution extending up to double the solar wind velocity has been detected in interplanetary space. This distribution unambiguously determines the source: interstellar neutrals, ionized and accelerated in the solar wind. The observed significant flux increase in early December is due to the gravitational focusing of the interstellar neutral wind on the downwind side of the Sun.

THE penetration of the interstellar medium into the heliosphere and the interaction between the solar wind and the interstellar gas have been of great interest for many years¹⁻⁴. The first experimental evidence of neutral interstellar hydrogen penetrating into the heliosphere was obtained from Lyman α sky background mapping^{5,6}. Similar observations of interstellar helium using the He I 584-Å resonance line⁷⁻⁹ revealed the existence of an interstellar neutral wind in interplanetary space which is subjected to the forces of solar gravitation and radiation pressure. When approaching the Sun, the interstellar gas is ionized by solar ultraviolet radiation, by charge exchange with solar wind ions and by collisions with solar wind electrons. The newly created ions are then picked up by the solar wind through interaction with the interplanetary magnetic field.

While most of the constituents are already ionized far beyond the orbit of the Earth, neutral helium (because of its high ionization potential) approaches the Sun to <1 AU (ref. 1). Therefore, a significant fraction of the helium is ionized inside the Earth's orbit and one would expect that these ions are observable by spacecraft in the solar wind. So far, no conclusive measurement on He^+ ions of interstellar origin has been given. Although signatures of He^+ in the solar wind with varying abundances have been reported occasionally^{10,11}, a systematic search for a permanently present flux of interstellar He^+ as part of the solar wind resulted only in upper limits more than one order of magnitude below the estimated flux levels¹². After this negative result Vasyliunas and Siscoe¹³ suggested that the newly created He^+ ions may not be thermalized rapidly enough to occur within the solar wind distribution. The expected broad energy and angular distribution of the interstellar helium ions, however, was not visible to the particle detectors flown on spacecraft up to now. A search for interplanetary He II emission at 304 Å was also not conclusive¹⁴. In this article, we will report measurements with a time-of-flight (TOF) spectrometer in which the interstellar helium ions and their velocity distribution in the solar wind are clearly identified for the first time.

Instrumentation and spacecraft

The data presented here were obtained with the SULEICA (Supra-thermal Energy Ionic Charge Analyser) instrument of the Max-Planck-Institut and the University of Maryland onboard the IRM spacecraft of the AMPTE (Active Magnetospheric Particle Tracer Explorer) project. The IRM was launched on 16 August 1984, into a highly elliptical orbit with an apogee of 18.6 R_E . During the time period from launch until December 1984 the satellite spent a large fraction of each orbit in the solar wind upstream of the Earth's bow shock.

The SULEICA instrument combines the selection of incoming ions according to their energy per charge by electrostatic deflection with a subsequent TOF analysis and the final measurement of the residual ion energy in a silicon surface barrier detector. The energy range from 5 to 270 keV/charge is covered by stepping the analyser voltage in 24 logarithmically spaced voltage steps in synchronism with the spacecraft spin. The fan-like

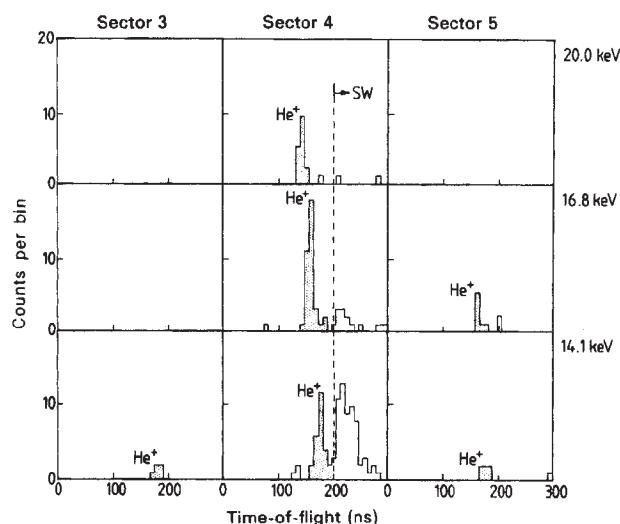


Fig. 1 Typical TOF-histograms at three different energy steps taken in the Sun sector and the two adjacent sectors. The data were obtained during a period of 45 min on 11 November 1984, at $\sim 18 R_E$ in front of the Earth's bow shock.

aperture of the electrostatic analyser (two concentric segments of a sphere) covers a solid angle of 10° in azimuth and 40° in elevation symmetric to the plane perpendicular to the spacecraft spin axis. The directional information in azimuth is provided by a sectoring scheme with 16 sectors for protons and α particles and eight sectors for the other ions. The geometrical factor of the instrument is $4.3 \times 10^{-2} \text{ cm}^2 \text{ sr}$ and its energy resolution is $\Delta E/E \approx 0.097$. A more detailed description may be found elsewhere¹⁵.

The expected energy of the pick-up ions (≤ 40 keV) is too low to create energy signals significantly above the noise level of the solid state detectors. Therefore, the ion species are identified by combining the electrostatic deflection (E/Q) and the TOF signal. For each specific E/Q step, the TOF distribution represents a unique mass-per-charge distribution, since $M/Q \sim E/Q \times \text{TOF}^2$. Recently this type of analysis has been successfully applied to identify Li^+ pick-up ions after the lithium vapour cloud releases of the AMPTE project in the solar wind¹⁶.

Basic observations

For the present investigation a limited number (12) of observational periods in the solar wind were chosen between launch of the spacecraft and the end of December 1984. During all periods we observe a significant peak at $M/Q = 4$ in the TOF-histograms for E/Q steps which correspond to energies higher than we would expect for genuine singly charged solar wind helium. Depending on the E/Q step we also observe a broad and variable peak at TOF-values which correspond to the actual

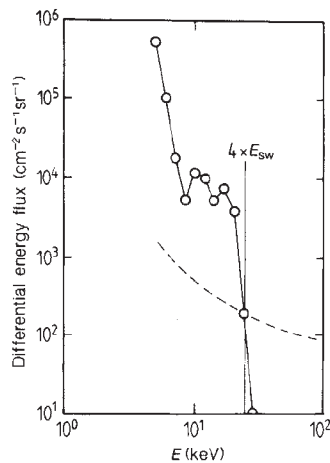


Fig. 2 Differential energy flux spectrum of the $M/Q = 4$ channel taken in the Sun sector with an accumulation time of 40 min. The dashed line represents the 1 count level for each energy.

solar wind velocity. These ions with M/Q values well above 4 represent heavy ions of the solar wind. In addition, the signature of the He^+ ions at $M/Q = 4$ is also seen in the two sectors adjacent to the sunward direction. A typical example of TOF-histograms for three different voltage steps and all three azimuthal sectors with the He^+ peak is shown in Fig. 1. The sample was taken during a 45-min interval on 11 November 1984, when no energetic ions occurred upstream of the bow shock. Note, however, that the He^+ signature is visible at all orientations of the interplanetary magnetic field and is independent of whether or not upstream particles are present.

Figure 2 shows the differential energy flux spectrum of the He^+ ions in the Sun sector for the same time period. The spectrum is basically flat between ≈ 8 and 22 keV with a sharp cutoff at ~ 22.3 keV. This value corresponds to about four times the solar wind bulk energy of helium (indicated by the vertical line) for this time period. The actual solar wind velocity was 550 km s^{-1} , as derived from the data of the Max-Planck-Institut/University of California Berkeley three-dimensional-Plasma instrument on IRM¹⁷. Below 8 keV a sharp rise of the spectrum occurs which is attributed to solar wind ions with $M/Q = 4$, for example, Si^{7+} and S^{8+} (ref. 12). The dashed line in Fig. 2 indicates the one count limit for He^+ .

Signature of pick-up ions

In contrast to genuine solar wind ions heavier than protons, which flow with almost solar wind bulk velocity, freshly created ions in interplanetary space are initially at rest. After ionization they are immediately subjected to the combined forces of the interplanetary $v_{\text{sw}} \times B$ electric field and the magnetic field B where v_{sw} is the solar wind velocity. In the inertial system (which coincides with the rest frame of the spacecraft), the ions initially perform a cycloidal motion perpendicular to the local magnetic field. The velocity varies between basically zero (the velocity of the neutral wind at $\approx 20 \text{ km s}^{-1}$ is neglected compared with the solar wind velocity) and a maximum value

$$v_{\perp \text{max}} = 2 v_{\text{sw}} \sin \alpha \quad (1)$$

which is determined by the solar wind velocity and the angle α between its flow direction and the local magnetic field. The maximum energy of pick-up ions then is

$$E_{\perp \text{max}} = 4(M/2)v_{\text{sw}}^2 \sin^2 \alpha \quad (2)$$

In the solar wind frame, the pitch angle distribution of these particles is a ring in velocity space with the pitch angle α , that is, the ions gyrate with $v_{\perp} = v_{\text{sw}} \sin \alpha$ and move along the magnetic field with $v_{\parallel} = v_{\text{sw}} \cos \alpha$. Observations of artificially-injected lithium in the solar wind which have been made within the first cycloid after the ionization are compatible with such

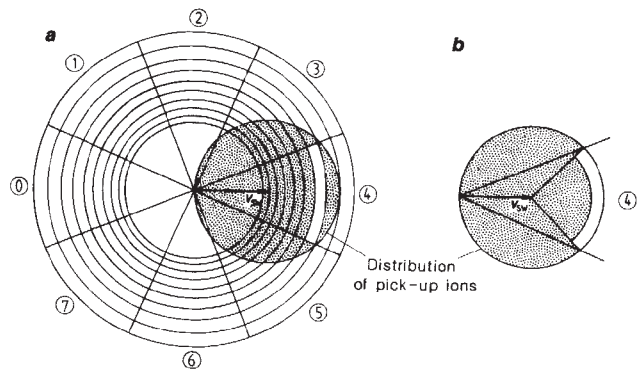


Fig. 3 *a*, Cut through the fully developed velocity distribution of pick-up ions in the spin plane of the spacecraft (shaded area) along with the sectoring and voltage stepping scheme of the instrument. The velocity space element, which is used to determine the differential energy flux as seen by the instrument, is left blank. *b*, Similar representation in the solar wind frame. The velocity space element, over which the actual integration is performed, is left blank.

an undisturbed pick-up distribution¹⁶. Our observations of He^+ ions, which are spread over a large volume in velocity space, seem to indicate a significant broadening of the initial ring distribution in pitch angle as well as in energy.

Cosmic rays are strongly pitch-angle-scattered by magnetic field fluctuations in the solar wind, while their energy remains almost unaffected¹⁸. The corresponding mean free path for scattering of He^+ pick-up ions (with a typical magnetic rigidity of 5–10 MV) turns out to be about 0.05 AU according to a compilation of cosmic-ray propagation parameters¹⁹. From these arguments, pitch angle scattering seems to be rather efficient for pick-up ions, and shortly after ionization a spherical shell distribution in velocity space is formed within the rest frame of the solar wind. Initially, the ions are injected into the solar wind with the energy $E_0 = (M/2)v_{\text{sw}}^2$ (with velocity components $v_{\parallel}$ and $v_{\perp}$ as discussed above). Therefore, the radius of the spherical shell distribution is equal to the solar wind velocity. Subsequently, the ion energy decreases as a result of adiabatic deceleration during the radial transport with the expanding solar wind. Due to adiabatic deceleration the energy of fully pitch-angle-scattered particles varies like

$$E/E_0 = \left(\frac{r_i}{r}\right)^{4/3} \quad (3)$$

with the distance r from the Sun, starting at the location of birth of the ion r_i (ref. 13). This relation immediately leads to a mapping of the spatial source distribution of pick-up ions upstream of the spacecraft into an energy or velocity distribution in the solar wind frame at the location of the satellite which can be written as:

$$f(v) = \frac{3N_0 \nu_{\text{ion}} r_0}{8\pi v_{\text{sw}}^4} \left(\frac{v}{v_{\text{sw}}}\right)^{-3/2} \quad (4)$$

N_0 and ν_{ion} are the neutral density and the ionization rate of the interstellar gas at the Earth's orbit ($r_0 = 1 \text{ AU}$). ν_{ion} is assumed to vary like $1/r^2$ with the intensity of solar ultraviolet radiation. In the spacecraft frame of reference, this results in a particle distribution which extends between $E = 0$ and a maximum energy

$$E_{\text{max}} = 4E_{\text{sw}} = 4(M/2)v_{\text{sw}}^2 \quad (5)$$

for particles coming from the sunward direction. Such a distribution with a more or less flat spectrum in differential energy flux is, indeed, observed between about 8 and 22 keV (see Fig. 2).

From these observations, it can be concluded that the velocity distribution of the pick-up ions is basically a full sphere centred at the solar wind velocity. A cut through this distribution in the plane perpendicular to the spin axis of the satellite is shown in Fig. 3. The azimuthal sectoring scheme and the voltage stepping

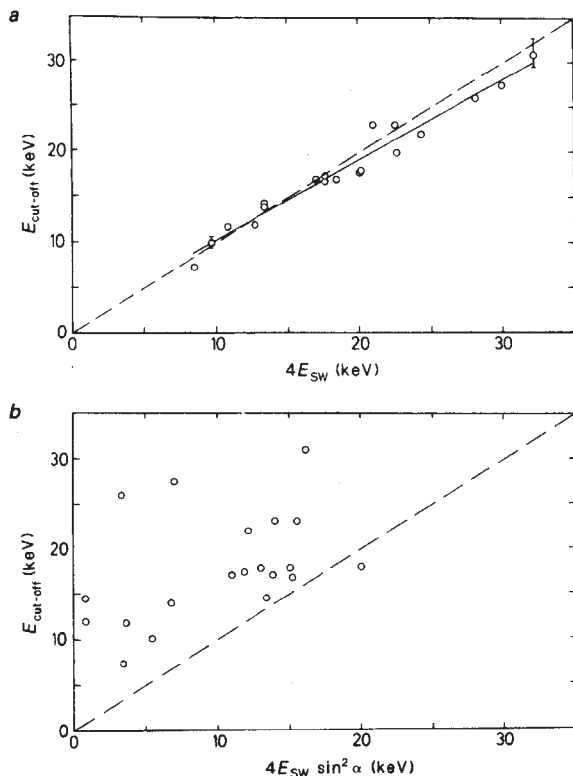


Fig. 4 Correlation of the experimental cutoff energies of the He⁺ spectrum with the fourfold solar wind bulk energy $4E_{SW}$ of helium (a) and with the maximum pick-up energy perpendicular to the interplanetary magnetic field $4E_{SW} \sin^2 \alpha$ (b).

scheme (concentric circles) are overlaid. This interpretation is also supported by the TOF histograms shown in Fig. 1, where the He⁺ peak is observed in the Sun sector 4 as well as in the neighbouring sectors 3 and 5. The energy spectrum displayed in Fig. 2 represents a cut through the distribution in the Sun sector. It follows from the measured flat energy spectrum that a large fraction of the sphere in velocity space rather than only the outer shell is filled with He⁺ ions. Otherwise, a significant flux decrease between E_{SW} and $4E_{SW}$ would be expected. Therefore, a significant part of the He⁺ population originates from locations several tenths of an AU upstream of the spacecraft. The energy of these ions has decreased on their way according to relation (3).

The energy spectra of the He⁺ ions have been investigated for several time periods of different solar wind velocities and interplanetary magnetic field directions. In Fig. 4, the observed cutoff energies E_{cutoff} (defined by a decrease of the differential energy flux by a factor of 10 compared with the mean plateau level) are plotted first against the fourfold solar wind bulk energy $4E_{SW}$ of helium as given by relation (5) and second against the initial maximum pick-up energy $4E_{SW} \sin^2 \alpha$ according to relation (2). There is an excellent correlation with the full four-fold solar wind energy (correlation coefficient 0.98), while the correlation with the maximum pick-up energy is very poor (correlation coefficient 0.36). This result clearly indicates that the pick-up ions in all conditions have mainly lost their directional information due to pitch angle scattering and that most of the ions originate from regions beyond one mean free scattering length (0.05 AU) upstream of the observer.

Note that the generation of several types of plasma waves by the highly unstable distribution of pick-up ions has been discussed in the past^{20,21}. These waves were expected to lead to a complete thermalization in energy of the pick-up ion distribution in the solar wind frame of reference in addition to pitch angle scattering. However, the results described above—in particular, the sharp cutoff in energy—indicate that the influence of these waves on the He⁺ distribution can be neglected compared with the adiabatic cooling.

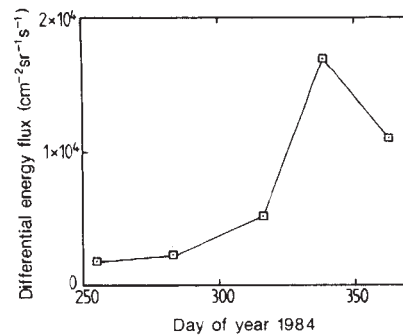


Fig. 5 Differential energy flux of 20-keV He⁺ ions obtained in the Sun sector for five different time periods between September and December 1984.

Source of the He⁺ ions

Knowing the velocity distribution of the He⁺ ions as shown in Fig. 3, the source strength $S = N_0 \nu_{ion}$ of the pick-up ions can be obtained from the measured differential energy flux at a fixed energy. The differential energy flux can be calculated by integrating the distribution function (relation (4)) over one voltage step and the Sun sector in the spacecraft frame (blank area in Fig. 3a). Instead of performing the full transformation from the solar wind frame into the spacecraft frame we use the fact that the distribution is isotropic in the solar wind frame and that near the maximum energy of the ion distribution (in the Sun sector) the volume element in velocity space $\Delta \Omega v^2 \Delta v$ covered by one energy channel of the instrument appears to be extended by a factor of four in solid angle in the solar wind frame compared with the spacecraft frame as indicated in Fig. 3b. Within this approximation, the integration is performed over the corresponding sector in the solar wind frame (blank area in Fig. 3b). Using the distribution function (5) we obtain for the differential energy flux:

$$\frac{\Delta J E}{\Delta E} = \frac{3 N_0 \nu_{ion} r_0 E}{8 \pi \Delta \Omega \Delta E} 4 \Delta \Omega \frac{1}{v_{sw}^4} \int_{0.9 v_{sw}}^{v_{sw}} \left(\frac{v}{v_{sw}} \right)^{-3/2} v^3 dv \quad (6)$$

According to the energy resolution of the instrument ($\Delta E/E \approx 0.1$, corresponding to $\Delta v/v \approx 0.05$) the integration limits in velocity are $0.9 v_{sw}$ and v_{sw} in the solar wind frame. Note that the differential energy flux (compared with the differential particle flux $\Delta J/\Delta E$), when taken at the maximum energy of the pick-up distribution, does not vary with the solar wind velocity. Therefore, this quantity is a suitable parameter for subsequent long-time studies of the He⁺ source. Using $w = v/v_{sw}$ we find:

$$\frac{\Delta J E}{\Delta E} = \frac{3}{2 \pi} N_0 \nu_{ion} r_0 \frac{E}{\Delta E} \int_{0.9}^1 w^{3/2} dw \approx 0.5 N_0 \nu_{ion} r_0 \quad (7)$$

Taking an ionization rate $\nu_{ion} = 1.25 \times 10^{-7} \text{ s}^{-1}$ at 1 AU (ref. 9) a value of $N_0 r_0 = 1.2 \times 10^{11} \text{ cm}^{-2}$ is derived from the mean differential energy flux between 16 and 20 keV ($7 \times 10^3 \text{ s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$) on 11 November 1984. The actual column density upstream of the spacecraft corresponds to about half the value of $N_0 r_0$.

Possible sources for singly charged ions which are injected into the solar wind with zero velocity are the interstellar neutral gas as discussed above and, in addition, planetary atmospheres and cometary comae. Of the latter ones the only source of interest for the observations reported here is the Earth's atmosphere. In contrast to the interstellar gas, which is present over a large scalelength in interplanetary space, the terrestrial sources of neutral helium are rather localized. Because the density of the exosphere typically varies like $1/R^2$ with distance from the Earth, the remaining scalelength of Earth's exosphere (the outer regime of the atmosphere) beyond the location of the spacecraft ($18 R_E$) is of the order of $20 R_E$. Therefore, a terrestrial origin can be ruled out for the observed He⁺ ions for two reasons.

First, the observed velocity distribution can only be explained in terms of a source which extends far beyond the mean free

scattering length of ~ 0.05 AU upstream from the Earth, while the terrestrial exosphere has a scalelength of typically $20 R_E$.

Second, with a scalelength of $\sim 20 R_E$ the measured column density would lead to a helium density of $\sim 5 \text{ cm}^{-3}$ at a distance of $18 R_E$. This value exceeds densities known for exospheric helium by more than 3 orders of magnitude²².

Therefore, the observed He^+ ions are most likely of interstellar origin. The corresponding helium density in interplanetary space upstream of the spacecraft on 11 November was determined to be $N_0 \approx 8 \times 10^{-3} \text{ cm}^{-3}$.

The flux of He^+ ions and, therefore, the neutral density of helium shows a significant variation over the observation period from September to December 1984. In Fig. 5, the differential energy flux of He^+ at 20 keV is plotted against time. For this analysis, only those time periods were taken when the solar wind velocity exceeded 550 km s^{-1} ; that is, the maximum energy E_{cutoff} was $> 20 \text{ keV}$. The differential energy flux is low in September and October ($\sim 2 \times 10^3 \text{ s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$). An increase by about a factor of 10 is observed until early December, after which the flux falls off again towards the end of December. This result is in good agreement with models of the spatial distribution of neutral helium in interplanetary space and the observations of $\text{He I } 584\text{-}\text{\AA}$ radiation^{8,9}. During December, the Earth is on the downwind side of the interstellar neutral wind with respect to the Sun. Here the helium is supposedly focused by the gravitational forces of the Sun and, therefore, the local density is enhanced significantly. The observation of this effect constitutes additional evidence for the interstellar source of the He^+ ions.

These observations are of particular importance, since various parameters of the local interstellar medium can be derived from the annual variation of the flux of He^+ ions: (1) the location of the apex; (2) the value of the relative velocity between the sun and the local interstellar medium; (3) the temperature; and (4) the density of the interstellar helium.

The absolute value of the density near the Earth as given above is compatible with results reported from ultraviolet measurements^{8,9}. A quantitative analysis of the density outside the heliosphere and the determination of the parameters of the local interstellar medium as listed under items (1) to (4) above requires a detailed comparison of the experimental results with

theoretical models of the neutral helium distribution in the heliosphere. This is beyond the scope of the present paper and will be presented in future work.

Furthermore, these pick-up ions represent a source of ions with a velocity distribution clearly distinguishable from the solar wind. To study their subsequent acceleration at the Earth's bow shock and at interplanetary shocks will be of great interest. Finally, these ions are most likely the seed particles for the so-called anomalous component of cosmic rays for which an interstellar origin and subsequent acceleration in the interplanetary medium²⁴ or at the terminating shock of the heliosphere²⁵ has been proposed.

We thank the many individuals at the Max-Planck-Institut and the University of Maryland who contributed to the success of the SULEICA instrument and to the AMPTE/IRM satellite. We thank H. Lühr for magnetic field data, G. Paschmann for solar wind velocity data, and M. A. Lee for helpful discussions in the interpretation of the ion distribution.

Received 22 July; accepted 23 September 1985.

1. Axford, W. I. *NASA SP-308*, 609-660 (1972).
2. Fahr, H. J. *Space Sci. Rev.* **15**, 483-540 (1974).
3. Holzer, T. E. *Rev. Geophys. Space Phys.* **15**, 467-490 (1977).
4. Thomas, G. E. A. *Rev. Earth planet. Sci.* **6**, 173-204 (1978).
5. Thomas, G. E. & Krassa, R. F. *Astr. Astrophys.* **11**, 218-233 (1971).
6. Bertaux, J. L. & Blamont, J. E. *Astr. Astrophys.* **11**, 200-217 (1971).
7. Paresce, F. & Bowyer, S. *Astr. Astrophys.* **27**, 399-406 (1973).
8. Weller, C. S. & Meier, R. R. *Astrophys. J.* **193**, 471-476 (1974).
9. Dalaudier, F., Bertaux, J. L., Kurt, V. G. & Mironova, E. N. *Astr. Astrophys.* **134**, 171-184 (1984).
10. Bame, S. J., Hundhausen, A. J., Asbridge, J. R. & Strong, I. B. *Phys. Rev. Lett.* **20**, 393-395 (1968).
11. Schwenn, R., Rosenbauer, H. & Mühlhäuser, K. H. *Geophys. Res. Lett.* **7**, 201-204 (1980).
12. Feldman, W. C., Asbridge, J. R. & Bame, S. J. *J. geophys. Res.* **79**, 1808-1812 (1974).
13. Vasyliunas, V. M. & Siscoe, G. L. *J. geophys. Res.* **81**, 1247-1252 (1976).
14. Paresce, F., Fahr, H. J. & Lay, G. J. *geophys. Res.* **86**, 10038-10048 (1983).
15. Möbius, E. *et al. IEEE Trans. Geosci. Remote Sensing GE-23*, 274-279 (1985).
16. Möbius, E. *et al. J. geophys. Res.* (in the press).
17. Paschmann, G. *et al. IEEE Trans. Geosci. Remote Sensing GE-23*, 262-266 (1985).
18. Jokipii, J. R. *Rev. Geophys. Space Phys.* **9**, 27-87 (1971).
19. Mason, G. M. *et al. Astrophys. J.* **267**, 844-867 (1983).
20. Wu, C. S. & Davidson, R. C. *J. geophys. Res.* **77**, 5399-5406 (1972).
21. Hartle, R. E. & Wu, C. S. *J. geophys. Res.* **78**, 5802-5807 (1973).
22. Paul, G. thesis, Univ. Bonn (1979).
23. Fahr, H. J. *Space Res.* **13**, 837-842 (1973).
24. Fisk, L. A. *Astrophys. J.* **206**, 333-341 (1976).
25. Pesses, M. *et al. Astrophys. J. Lett.* **246**, L85-L88 (1981).

Axial processes along a segment of the East Pacific Rise, 10° – 12° N

G. Thompson*, W. B. Bryan*, R. Ballard*, K. Hamuro† & W. G. Melson‡

* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

† Japanese Consulate, 600 Atlantic Ave, Boston, Massachusetts, USA

‡ Smithsonian Institution, Washington DC 20550, USA

Axial segments of the East Pacific Rise are made up of individual volcanoes. Each volcano has a distinct magma composition and shows a systematic variation in the fractional crystallization with distance along the axis from the central chamber. Hydrothermal venting, lava morphology and tectonics also vary along the axis. Lavas erupted near transform faults, from the tips of propagating rifts in overlapping spreading centres, and from near-axis seamounts have different and variable compositions.

MID-OCEAN ridge spreading centres have recently been shown to consist of discrete segments separated by transform faults and conjugate spreading centres (also called overlapping spreading centres (OSC) or non-transform offsets^{1,2}). Typically, in each segment the axial depth shows a minimum or topographic high at some point between the two ends¹⁻⁴. Francheteau and Ballard⁴ have proposed that the topographic highs lie above the principal magma chambers feeding lavas to the axis and are thermal expressions of such chambers. Their model predicts that relative volumes of lava and morphology of lava flows

should change systematically along the axis away from the topographic high. Specifically, pillow lavas should increase relative to sheet flows, fracturing should increase and hydrothermal activity should decrease along-ridge away from the topographic high. To date, observations supporting this hypothesis have been made at several locations of different spreading rates, but have been limited to a few kilometres along the axis in the vicinity of the axial topographic high.

We report here on our recent observations over a 180-km length of East Pacific Rise (EPR) axis between 10° and 12° N,

made specifically to test the Francheteau-Ballard model. This location (see Fig. 1) was deliberately chosen because the region had been recently mapped using SEABEAM⁵ and the axial depth and segmentation of the ridge was known. Starting at a major transform fault, the Clipperton Fracture Zone, the neovolcanic zone or zero-age axial depth shows a progressive shallowing to a first topographic high about 10°55' N, deepens to a topographic low, thought to be a small OSC at 11°15' N (ref. 1), rises to a second topographic high at 11°30' N, then deepens to a large OSC at 11°45' N with a distinct basin separating the overlapping ridge segments (Fig. 1). Previously, a French submersible dive⁶⁻⁸ on the second topographic high at 11°30' N had indicated intense (black smoker) hydrothermal activity. Thus, all the ingredients were present to test how axial processes varied along the axis with respect to major features such as topographic highs, OSCs and transform faults, and how (or if) magmatic activity and composition varied with distance from the proposed central magma chamber(s).

Variation along axis

Near the Clipperton Fracture Zone at the southern end of the studied ridge segment, the axial graben is not well developed, abundant fissures and talus indicate dominance of tectonic activity, and pillow lavas are the major eruptive form (but are often badly cracked; see Fig. 2a). These observations indicate that eruptions are infrequent. No hydrothermal activity was observed. Halfway along the segment between the Clipperton Fracture Zone and the first topographic high, the axial graben becomes well defined, although fissures are abundant near the graben walls. Within the graben, glassy lobate flows and broken platy sheet-flow lavas (Fig. 2b) indicate relatively recent eruptions, although tectonic activity is still noticeable. Some non-active hydrothermal formations were observed, although these are rare (Fig. 2c). At the first topographic high at 10°55' N (see Fig. 1b), the axial graben is very well developed, undeformed sheet-flow lavas (Fig. 2d) dominate in the graben floor, and hydrothermal activity, including white and black 'smokers', is common (Fig. 2e). Fissuring and faulting near the graben edges are not as abundant as farther south in the segment. Constructive volcanism dominates the topography. Glassy reflective flows are not common in the photographs but sampling recovered very fresh, glassy fragments of sheet flows and lobate flows with a thin mantling of iron-rich hydrothermal precipitate that dulls the surfaces.

The topographic low between the two highs, although previously indicated¹ to be an OSC, does not appear to have typical OSC features. The axial graben appears to be continuous and fresh lava flows and hydrothermal activity are seen. We infer, based on the observations and the magma chemistry, that the axial portion from the northernmost topographic high at 11°30' N has propagated southwards and linked up with the southern axis. The second topographic high at 11°30' N (see Fig. 1b) also shows a well-developed axial graben, sheet flows dominate, and hydrothermal activity with concomitant vent communities and biota are abundant. A cross-section across the graben here, as at the first topographic high, shows only pillow lavas outside the graben, fissuring near the graben wall, talus, collapse pits and flow-back structures in the graben close to the wall, and sheet flows and lava 'ponds' flooring the graben. Hydrothermal activity is found only within or at the walls of the graben⁹. These observations indicate that the sequence of eruptions in the graben must be sheet flows, followed by lobate flows, and finally by pillow lavas as the graben fills. Eventually, the ridge splits, a new graben is formed, and the cycle repeats.

The axial region north of the second topographic high at 11°30' N shows a marked transition as it becomes part of the OSC. Tectonically undisturbed lava formations are less abundant, pillow lavas and talus dominate, the axial graben shallows or pinches out, and hydrothermal activity is not seen. This contrasts markedly with the eastern limb of this OSC which shows active hydrothermal vents, recent sheet flows, and a well developed graben. We believe this limb is active and is propagat-

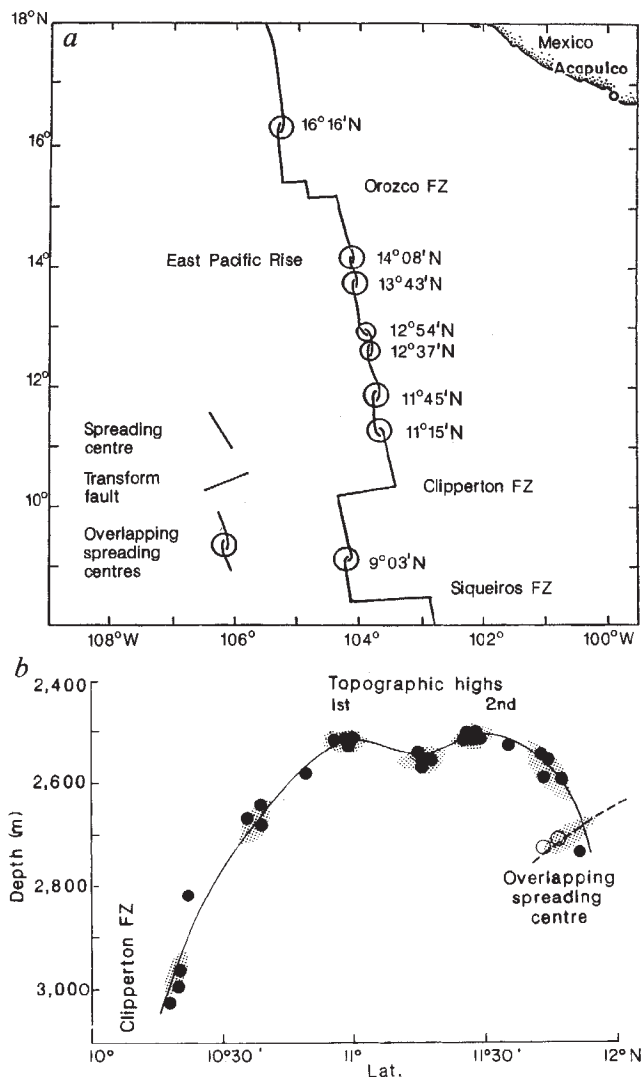


Fig. 1 a, EPR axis showing location of transform faults and conjugate (overlapping) spreading centres. The region studied in this work is the EPR axis from the Clipperton Fracture Zone to, and including, the OSC at 12° N. b, Axial depth versus latitude for the EPR 10°-12° N. The hatched areas are regions photographed using transponder navigation. ●, Dredge locations; ○, dredge locations in the east limb of the OSC. We placed transponders at seven discrete localities along the 180 km of EPR axis and performed detailed ANGUS¹⁴ photography over 4-7 km of the central axial region (see Fig. 1b for location of nets). This photographic coverage allowed us to delineate the axial graben and document such features as lava morphology, relative ages of eruptions, tectonic activity and hydrothermal activity. Using the transponder nets and information from the bottom photographs, we did carefully navigated dredging within the nets and within the neovolcanic zone, plus some dredging between nets using the SEABEAM bathymetry and conventional 12-kHz echo-sounding to guide the sampling (see Fig. 1b for sample location). In all, we made 14 ANGUS runs and 23 successful dredge hauls. About 70 samples initially selected as representative of type from each locality were analysed for major and trace elements by X-ray fluorescence spectroscopy at Woods Hole¹⁵, glass margins were analysed using an electron microprobe at the Smithsonian Institution¹⁶, and petrographic studies and mineral analyses were made on selected thin sections. These preliminary data are presented here.

ing southwards. The nodal basin between the overlapping limbs appears to be a passive feature. The walls are composed only of lava spillage or talus from the respective axes and the floor is covered with sediment, with only a few, old pillow lavas showing through the sediment (Fig. 2f). There appears to be no current volcanic activity within the OSC basin, and no tectonic activity or faulting suggestive of tensional or transform move-

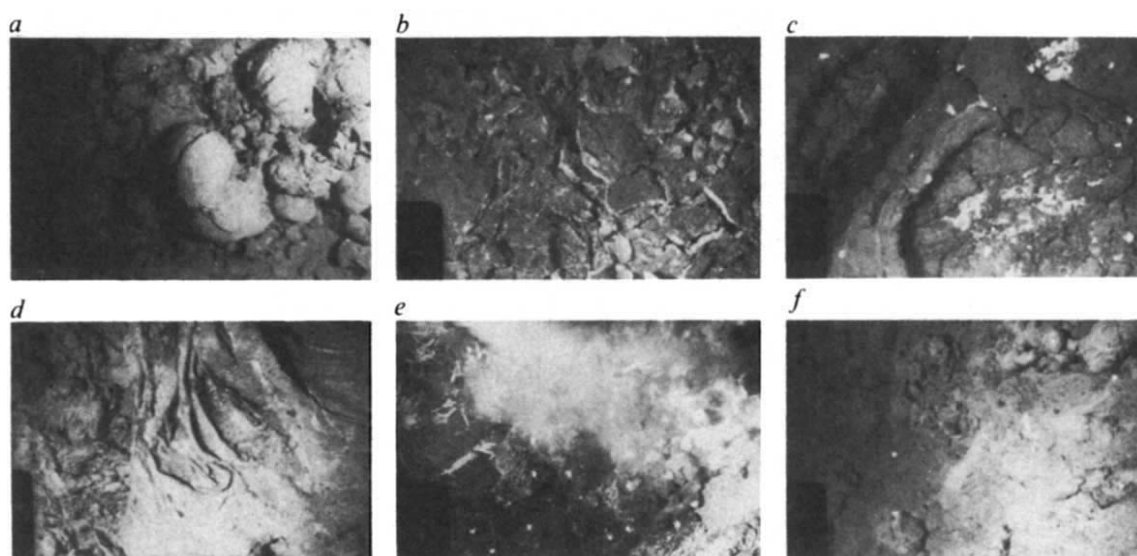


Fig. 2 *a*, EPR axis near the Clipperton Fracture Zone showing typical pillow formations, often cracked. *b*, EPR axis a few kilometres north of *a* location, showing cracked and broken sheet-flow plates in axial graben. *c*, EPR axis about halfway between the Clipperton Fracture Zone and the first topographic high, showing an extinct hydrothermal vent with dissolving clam shells in collapsed lobate flow in the axial graben. *d*, EPR axis at first topographic high showing well developed and fresh sheet-flow formations on floor of axial graben. *e*, Hydrothermal 'smoker' near wall of axial graben on topographic high. Such vents are common on both axial topographic highs. *f*, Floor of nodal basin between the two axial limbs of the OSC at 12° N. The floor is undisturbed and sediment-covered with a few old pillow-lava formations showing.

ment between the limbs as previously suggested for such features¹⁰.

Various petrographic and chemical discriminants allow us to define five magma groups, A-E (Fig. 3). The lavas of magma groups B and C along this segment of the EPR show a systematic variation in Mg number ($\text{Mg}/(\text{Mg} + \text{Fe}) \times 100$) (Fig. 3a). With the exception of a few relatively 'primitive' members of magma group A, eruptives farthest from the topographic highs have lowest Mg numbers, suggesting possible cooling and fractionation as magma migrates along-rift. Group B basalts at the low point near the Clipperton Fracture Zone are apparently most fractionated, and there is a systematic decrease in the implied degree of fractionation approaching each topographic high. Basalts increase again in apparent degree of fractionation as the axis deepens into the large northernmost OSC. Both limbs of the OSC, including the tip of the propagating eastern limb, have these apparently fractionated basalts. The systematic compositional variation is reflected in covariances of most chemical components and by the modal mineralogy and mineral compositions. The most fractionated group B varieties, as inferred from Mg number, are moderately phryic basalt with up to 25% plagioclase, pyroxene, and minor olivine phenocrysts. Approaching the topographic highs the basalts become less phryic, only plagioclase and minor pyroxene or olivine are present, and An and Fo contents increase. Around the first topographic high the basalts are aphyric, and only minor plagioclase microphenocrysts are present. From 11°15' N to the second high, plagioclase plus olivine microphenocrysts are common, while pyroxene is present only near 11°15' N. The most phryic group C basalts (~15% plagioclase, olivine and pyroxene) are at the north end of the rift, in the overlap zone at about 11°50' N. Preliminary calculations indicate that on the first topographic high and along the southern limb to the Clipperton Fracture Zone, with the exception of a few basalts nearest the fracture zone, the basalts are comagmatic and simple low-pressure fractionation can be modelled using observed mineral and magma compositions. Similar relations have been demonstrated within the second topographic high and its axial limbs (W.B.B. and K.H., in preparation).

Some of the differences in both major- and trace-element chemistry of the five different magma types can be seen in Fig. 3. A plot of FeO versus MgO (Fig. 3b) shows consistent but subtle

differences between groups B and C, while the group of lavas near the Clipperton Fracture Zone (magma type A), and the magmas in the eastern limb of the OSC (magma type D) are clearly defined. Magma group E is found at two locations on the northernmost topographic high; it is characterized by large enrichments in incompatible elements compared with the other basalts from the EPR axis and is similar to transitional or enriched mid-ocean ridge basalts (E-MORB). The most distinctive chemical features of these five magma types are shown in Table 1.

Basalts of magma type A near the Clipperton Fracture Zone occur in the vicinity (same dredge haul) of other basalts belonging to magma type B. However, as noted in the photographic evidence, the axial region is badly broken up and different eruptive units and ages are in close proximity in fault scarps and talus. The basalts of magma type A show a wide range in degree of fractionation compared with the adjacent members of type B magma, which are all relatively fractionated. Compared with magma B, magma A shows lower Fe and higher Ti content for a given MgO concentration. Trace-element contents also are distinctive compared with magma type B, with lower Sr contents (see Fig. 3c), lower Zr/Y, and higher Zr/Nb ratios (see Fig. 3d). Langmuir and Bender¹¹ have previously noted that basalts erupted close to major transforms have different and unique compositions compared with adjacent axial basalts

Table 1 Chemical characteristics of different magma types along EPR axis 10°–12° N

	Type A	Type B	Type C	Type D	Type E
MgO (wt %)	5.5–7.4	5.6–7.2	6.4–8.1	6.1–6.8	6.6–7.4
FeO* (wt %)	10.9–13.2	10.4–13.8	9.2–11.9	12.2–13.7	9.4–10.0
TiO ₂ (wt %)	1.6–2.4	1.7–2.6	1.5–2.2	2.0–2.4	2.1–2.3
K ₂ O (wt %)	0.14–0.25	0.12–0.25	0.11–0.22	0.09–0.12	0.62–0.73
Sr (p.p.m.)	78–122	129–140	141–168	93–100	196–320
Zr/Nb	39–52	25–35	20–30	38–44	8.7–10.2
Zr/Y	2.4–3.2	2.9–3.0	3.0–3.3	2.5–2.7	4.1–4.6

* Total Fe as FeO.

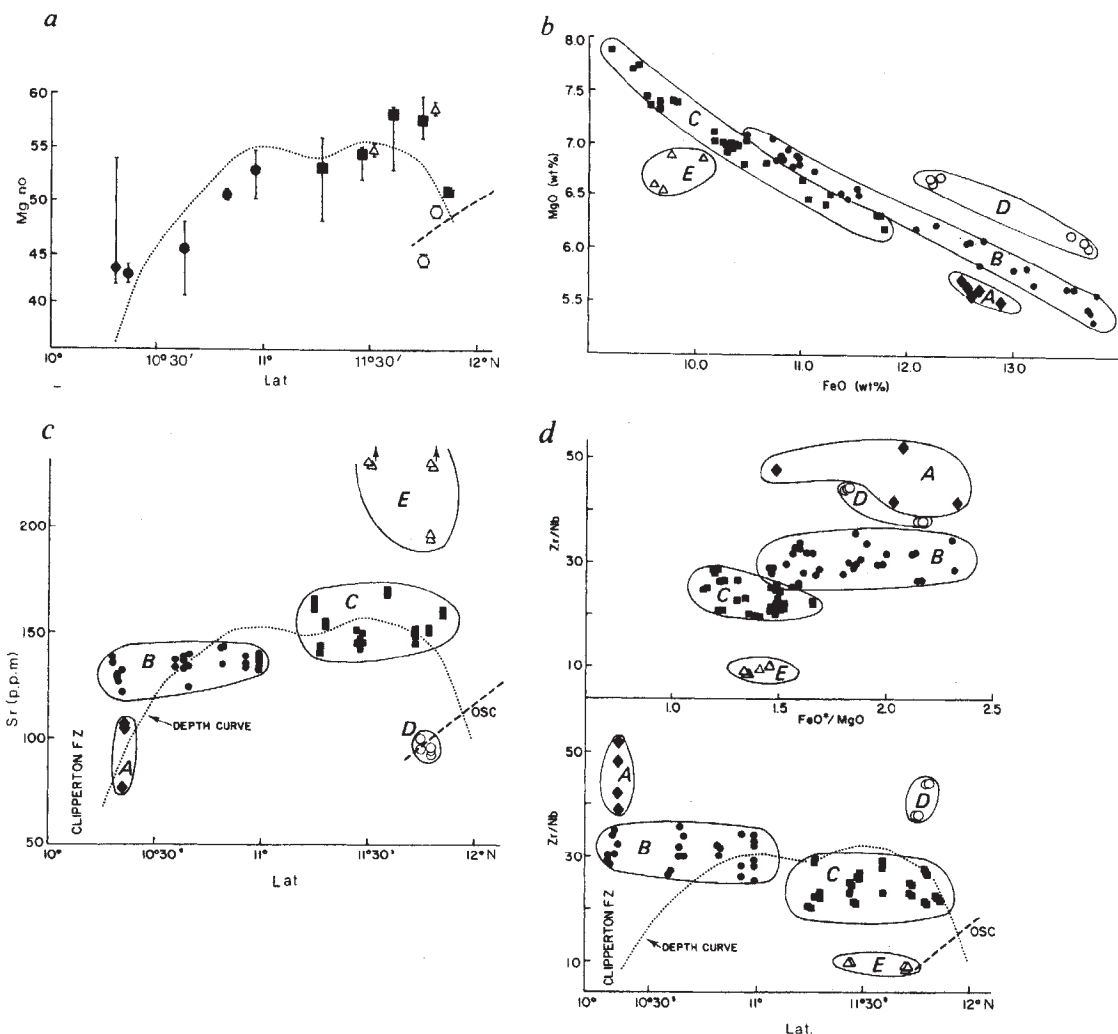


Fig. 3 *a*, Mg number ($\text{Mg}/(\text{Mg} + \text{Fe}) \times 100$) versus latitude along EPR axis 10° – 12° N. Symbols are mean Mg numbers, and the vertical line shows the range of Mg numbers at that location. *b*, MgO versus FeO for basalt glasses from the EPR axis. Five different groupings are noted as well as the range of FeO/MgO, indicating the degree of fractionation. *c*, Sr versus latitude for basalts from EPR axis. Five distinct magma types and their location are distinguished. *d*, Zr/Nb versus FeO*/MgO and latitude for basalts from EPR axis. ♦, A, near Clipperton Fracture Zone; ●, B, South axis and topographic high; ■, C, north axis, topographic high and west OSC; ○, D, eastern OSC; △, E, E-MORB. The dotted line in *a*, *b*, and *d* is the axial depth (see Fig. 1b).

and have referred to this as the “transform edge effect” (TFE). We are unsure without further analysis (such as isotopes, rare-earth elements, Ta, Th, and so on) whether magma type A is another example of a TFE or merely an older axial eruptive event unrelated to the present axial eruptives represented by magma type B.

Magma type B shows a wide range in degree of fractionation which decreases systematically towards the topographic high. This is in agreement with the model suggesting major input to a master magma chamber beneath the topographic high. Lateral injection along-ridge (probably sub-crustally) from the chamber with concomitant crystallization as the magma moves into cooler crust, is a reasonable explanation which is also consistent with recent observations in Iceland¹². All the basalts from magma type B are normal MORB, as are those from magma type C from the second topographic high. Single specimens from either magma probably could not be discriminated as different types, but the large number of samples analysed here with high precision show small but consistent differences in both chemistry and modal phase proportions between the two main magma types B and C which persist across substantial variations in fractionation. Mineralogically, magma type B has plagioclase and pyroxene with little olivine, whereas magma type C has both olivine and plagioclase as phenocrysts but little pyroxene. Magma type B generally has higher Fe and lower Ti content for a given MgO concentration, and also has lower Sr contents,

lower Zr/Y ratios and higher Zr/Nb ratios compared with magma C (Fig. 3b–d). The differences between magma B and C are small and compositional ranges overlap. However, comparisons of trace-element ratios versus Mg number or FeO/MgO ratio show that the magmas are distinctly different in composition.

The topographic low between the axial highs is apparently filled with type C lavas, suggesting that the southward propagating limb of the northern high has almost closed this small OSC. The northernmost limb of the rift containing magma type C also becomes increasingly fractionated but as the distances are not so great, and the tip of this limb is apparently not active, the degree of fractionation is not so pronounced as on the long southern limb of the first topographic high of magma type B.

The magmas from the active eastern limb of the OSC, magma type D, have quite unique compositions compared with the main lavas of types B and C; they resemble group A in some but not all features. They have higher Fe and Ti contents for a given MgO concentration, low K and Sr content, low Zr/Y and high Zr/Nb ratios. They are phryic basalts in which olivine is a prominent phenocryst and include some varieties which are quite fractionated, but we are unsure whether they are fractionated extremities of lavas from the main chamber near $12^{\circ}30'$ N, or lavas showing a TFE effect since they are at the tip of a propagating rift breaking into old, colder crust¹³.

Magma type E lavas, the so called E-MORB, were recovered

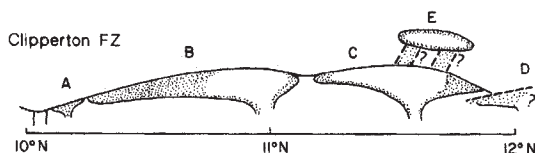


Fig. 4 Five magma types observed along EPR axis 10°–12° N. Dotted areas indicate >8% mineral phenocrysts. Magma type E is presumed to be an off-axis (4 km) seamount from which lavas have intruded into the axial region at two locations. It is uncertain whether magma type D is a unique magma batch like magma type A, near the propagating rift tip, or the proximal tip of a down-axis injection from a magma chamber near 12°30' at the next axial topographic high.

at two localities on the second topographic high. A previous cruise by the French *N.O.J. Charcot*⁷ in the same general area also recovered samples of E-MORB from the axial graben and from the eastern rim of the axis near 11°27' N. We recovered two samples about 20 km away in the northern part of the axis approaching the tip of the west limb of the OSC. R. Hekinian provided splits of the French samples which we analysed; these are E-MORBs similar to those we recovered. The type E lavas have higher Ti and lower Fe content for a given MgO concentration compared with the other axial lavas. They have very high concentrations of incompatible elements such as K, Sr, Zr and Nb, and higher Zr/Y and lower Zr/Nb ratios than the other magma types. The *Charcot* surveys have shown that at this latitude, about 11°45' N, there is a large seamount only about 4 km to the west of the ridge axis which is erupting various types of lavas (R. Hekinian, personal communication). Available data are not sufficient to determine whether lavas from this seamount may have been injected transversely, or whether they flowed on the sea floor from the seamount to the present EPR axis.

Conclusion

Figure 4 shows diagrammatically the distribution of the five magma types. Our observations support the Francheteau-Ballard model of axial processes. We suggest that the topographic highs are the locus of magma upwelling into magma

chambers, and that magmas are injected down the axes from these localities. These lavas show systematic changes in composition along the axis. In essence, the EPR is made up of a series of volcanoes separated by transform or non-transform (OSC) offsets. This is analogous to the observations of central volcanoes located along the rift zone in Iceland which have been shown to inject magma, subcrustally, up to 70 km along rifts from the central volcano; this injection often is highly asymmetric as in these EPR examples¹².

We have shown that the depression between the overlapping limbs of an OSC is old and inactive. We have demonstrated that the eastern limb of the OSC at 11°45' N is very active and propagating southward. Our data indicate that lavas at the propagating tip of an overlapping pair of rifts, or close to a large transform, may have distinct chemical compositions. Our observations also suggest local chemical heterogeneity in the mantle that must give rise to the pronounced differences between axial and nearby seamount lavas. We also observe small-scale mantle heterogeneity or differences in melt formation, local plumbing or magma mixing that give rise to the distinct compositions observed between individual magma centres or volcanoes along axis. These differences are superimposed on compositional variation which can be produced by the fairly extensive fractionation which has occurred within each magma batch. Further aspects of these differences and the nature of the source regions are currently being investigated by isotope and rare-earth element analysis and geochemical modelling. Although a variety of magma compositions have been noted in samples recovered previously from the EPR, this is the first documentation of systematic variation in lava composition and morphology that can be ascribed to specific axial processes and their relative position in time and space.

We thank the officers, crew and technicians of R/V *Melville* for assistance, and Rob Zierenberg (USGS) for help in dredging. We thank Brian Schroeder, Margaret Sulanowska and Tim O'Hearn for help in the analysis of the samples. J. Marsh and M. Harvey typed the manuscript. The manuscript benefited from reviews by R. Hekinian. This work was supported by NSF grant OCE83-09977. This is contribution no. 5962 from the Woods Hole Oceanographic Institution.

Received 21 June; accepted 30 September 1985.

1. Macdonald, K. & Fox, P. J. *Nature* **302**, 55–58 (1983).
2. Schouten, H. & Klitgord, K. D. *Nature* **303**, 549–550 (1983).
3. Lonsdale, P. *Bull. geol. Soc. Am.* **96**, 313–327 (1985).
4. Francheteau, J. & Ballard, R. D. *Earth planet. Sci. Lett.* **64**, 93–116 (1983).
5. Macdonald, K., Sempere, J. C. & Fox, P. J. *J. geophys. Res.* **89**, 6049–6069 (1984).
6. Ballard, R. D., Hekinian, R. & Francheteau, J. *Earth planet. Sci. Lett.* **69**, 176–186 (1984).
7. Hekinian, R., Francheteau, J. & Ballard, R. D. *Oceanol. acta* **8**, 147–155 (1985).
8. Hekinian, R. *et al. Science* **219**, 1321–1324 (1983).
9. McConachy, T. F., Mottl, M. J. & Von Herzen, R. P. *EOS* **65**, 1124 (1984).

10. Lonsdale, P. J. *geophys. Res.* **88**, 9393–9406 (1983).
11. Langmuir, C. H. & Bender, J. F. *Earth planet. Sci. Lett.* **69**, 107–127 (1984).
12. Sigurdsson, H. & Sparks, S. R. J. *Nature* **274**, 126–130 (1978).
13. Sinton, J. M., Wilson, D. S., Christie, D. M., Hey, R. N. & Delaney, J. R. *Earth planet. Sci. Lett.* **62**, 193–207 (1983).
14. Phillips, J. D., Driscoll, A. H. J., Peal, K. R., Marquet, W. M. & Owen, D. M. *Deep-Sea Res.* **26**, 211–225 (1979).
15. Schroeder, B., Thompson, G., Sulanowska, M. & Ludden, J. N. *X-ray Spectrom.* **9**, 198–205 (1980).
16. Melson, W. G. *et al. Smithson. Contr. Earth Sci.* **19**, 31–60 (1977).

Patterns of engrailed and fushi tarazu transcripts reveal novel intermediate stages in *Drosophila* segmentation

Michael P. Weir & Thomas Kornberg

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA

Transcripts of the engrailed and fushi tarazu genes in young Drosophila embryos are initially patterned in intervals larger than a single segment. The segmental pattern evolves in a stepwise sequence through successively smaller spatial units.

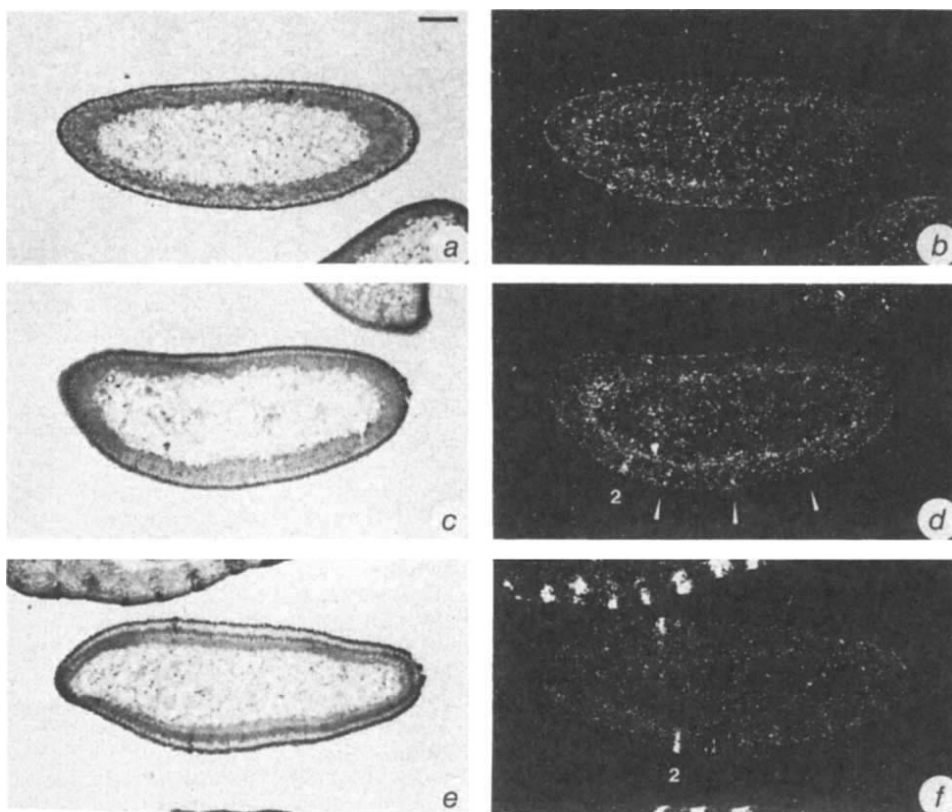
THE *Drosophila* embryo is spatially organized into repeating units—the anterior and posterior developmental compartments¹ of each metameric segment. These compartment types alternate in a striking zebra-like pattern as transverse encircling bands of cells on the surface of the young embryo^{2,3}. Their presence is

not apparent from the embryo morphology, but is revealed by cell lineage studies^{4–6}, and by *in situ* identification of cells that express the engrailed locus. The engrailed locus is expressed only in the posterior compartment cells².

The engrailed gene is one of a number of genes whose function

Fig. 1 Micrographs showing *engrailed* transcripts in blastoderm stage embryos. Bright-field (left) and corresponding dark-field (right) micrographs were recorded after hybridization with an *engrailed* cDNA probe and autoradiography. Sections in all figures are oriented anterior left. *a, b*, Section of embryo at nuclear cycle 13 with no detectable localization of autoradiographic grains. *c, d*, Early cycle 14 embryo, oriented ventral down, with nuclei beginning to elongate. Ten such sections were found with grains most prominently over the nuclei that later will correspond to the second of the 14 stripes (Fig. 2*f*). In this section, sparse concentrations of grains are also over the nuclei of what are probably stripes 4, 8 and 12 (arrowheads). *e, f*, Horizontal section. The nuclei have elongated and membrane furrows have extended into the cytoplasm just beyond the nuclei. A prominent autoradiographic signal is over the nuclei and cytoplasm of stripe 2. Scale bar, 50 μ m.

Methods. Embryos were dechorionated with 2.5% sodium hypochlorite for 3 min, fixed in 2% paraformaldehyde, 0.5 $\times$ phosphate-buffered saline (PBS) and 50% *n*-heptane for 15 min using vigorous shaking, and transferred to 45% methanol, 25 mM EGTA, 50% heptane and shaken on dry ice for 10 min. Vitelline membranes were then removed by rapid heating to 40 $^{\circ}$ C²⁷. Following three washes in 90% methanol, 50 mM EGTA, embryos were transferred in stages to PBS, before embedding in O.C.T. (Miles Scientific) using liquid nitrogen. *In situ* hybridizations onto 8 μ m frozen sections were performed as described previously using tritiated nick-translated DNA probes of the p9-1.4 cDNA². Embryonic stages: Observation during nuclear cycles 10-13, syncytial blastoderm stages, shows that at cycle 10, yolk nuclei stop dividing, pole cells bud from posterior end, and remaining nuclei appear on the surface and divide in near synchrony. At cycle 14, cellular blastoderm stage, there is a ~90 min period of mitotic quiescence during which membranes form between the peripheral nuclei and gastrulation begins. Embryos in the sections were staged by counting nuclei in 100- μ m intervals along the embryo periphery. 43 measurements of cycle 14 embryos yielded a value of 15.6 ± 1.4 nuclei per 100 μ m. The values used for other stages were: cycle 10, 4 nuclei; cycle 11, 6 nuclei; cycle 12, 8 nuclei; and cycle 13, 11 nuclei.



is essential for segmentation. The gene *fushi tarazu* (*ftz*) is another⁷, and it is expressed in the young embryo in seven evenly spaced bands of cells with a two-segment periodicity⁸. Although each band has the width of one segment⁹, the spatial relationship between the bands of *ftz*-expressing cells and the metamereric segments is not known. Nevertheless, the expression of *engrailed* and *ftz* both provide indications of the timing and nature of the processes that subdivide the embryo into equal-sized metameres, and they provide molecular markers that can be used to follow the process of segmentation.

In this study, the generation of the *engrailed* and *ftz* patterns of expression have been determined by *in situ* hybridization. The use of fixation methods that better preserve the morphology of the early *Drosophila* embryo and longer autoradiographic exposures have yielded a more precise portrayal of the generation of the segmental patterns and have revealed several new aspects of segmentation. Transcription patterns for both *engrailed* and *ftz* suggest that subdivision of the embryo proceeds through a sequence of intermediate stages, each stage part of a process that partitions the embryo into progressively smaller units.

Expression of *engrailed* in early embryos

After fertilization and nine synchronous mitotic divisions, three types of nuclei can be distinguished in the *Drosophila* (cycle 10) embryo: approximately 100 centrally located yolk nuclei, about 450 peripheral nuclei at the embryo surface that will generate the somatic tissues, and about 12 future germ cells at the posterior pole¹⁰⁻¹⁴. Four divisions of the peripheral nuclei follow. During the 14th cell cycle, the nuclei, initially round,

elongate before membrane furrows invaginate from the surface of the embryo to separate the nuclei into individual cells¹⁵. During this period the nuclei become highly active in transcription^{16,17} (A description of relevant embryonic stages is given in Fig. 1 legend.)

To identify nuclei that express the *engrailed* locus, wild-type embryos were fixed and quick-frozen for preparation of cryostat sections. After hybridization with radiolabelled probes generated from an *engrailed* cDNA clone², autoradiographs were exposed for 3-4 months to detect *engrailed* transcripts. Sections from embryos younger than stage 14 did not have grains localized either over nuclei or in any other distinctive pattern (Fig. 1*a, b*), yet the grain density over such embryos was consistently higher than background. Since Northern blot analysis of RNA isolated from cleavage stage and syncytial blastoderm stage embryos has revealed the presence of *engrailed* transcripts¹⁸, these early transcripts are either not regionally localized or our *in situ* hybridization methods lacked sufficient sensitivity for definitive detection.

As the embryo matures through the 14th division cycle, striking changes in *engrailed* expression produce a pattern of 14 equally spaced and equally dense stripes (Fig. 2*f*). The pattern is generated in a reproducible and recognizable sequence. Anterior bands appear before posterior ones, ventral bands before dorsal ones (Figs 1, 2). However, superimposed on this anterior to posterior sequence of appearance are strong but transient periodicities in the density of autoradiographic grains over the different bands: (1) the first two prominent bands in the ventral region are bands 2 and 8; (2) bands 4 and 8 become prominent before band 6, and band 12 before 10, suggesting a

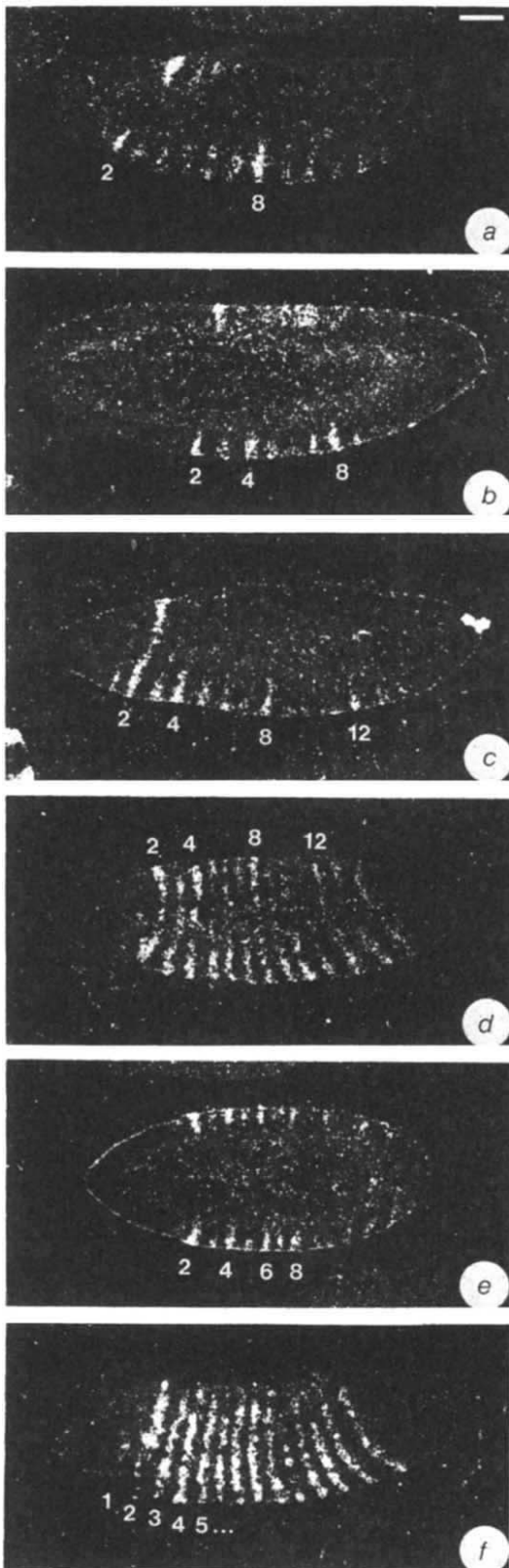


Fig. 2 Micrographs of *engrailed* transcripts in early gastrulas. Dark-field micrographs of embryos oriented ventral down except for *e*, a horizontal section. Stripes are more prominent in the anterior and ventral portions. Most prominent stripes are: *a*, ventral 2 and 8, dorsal 2; *b*, ventral 2, 4 and 8, dorsal 2; *c*, ventral 2, 4, 8 and 12, dorsal 2; *d*, ventral 2-14 and dorsal 2, 4, 8 and 12; *e*, lateral 2, 4, 6, 8, 10, 12; *f*, all stripes approximately equal. Note that among the weak stripes between 4 and 8 in *b*, stripe 6 is the weakest. *a* Has a two-band pattern, *b-d*, four-segment periodicity, and *e*, a two-segment periodicity. Scale bar 50 μ m.

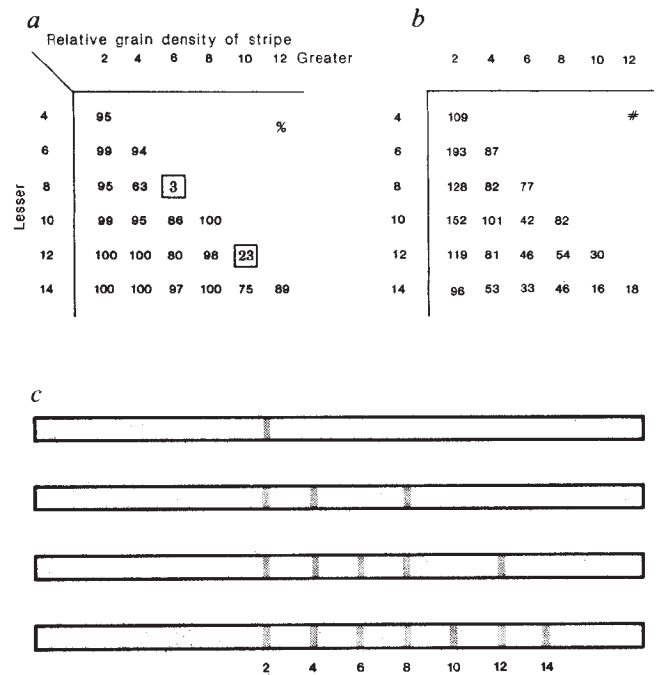


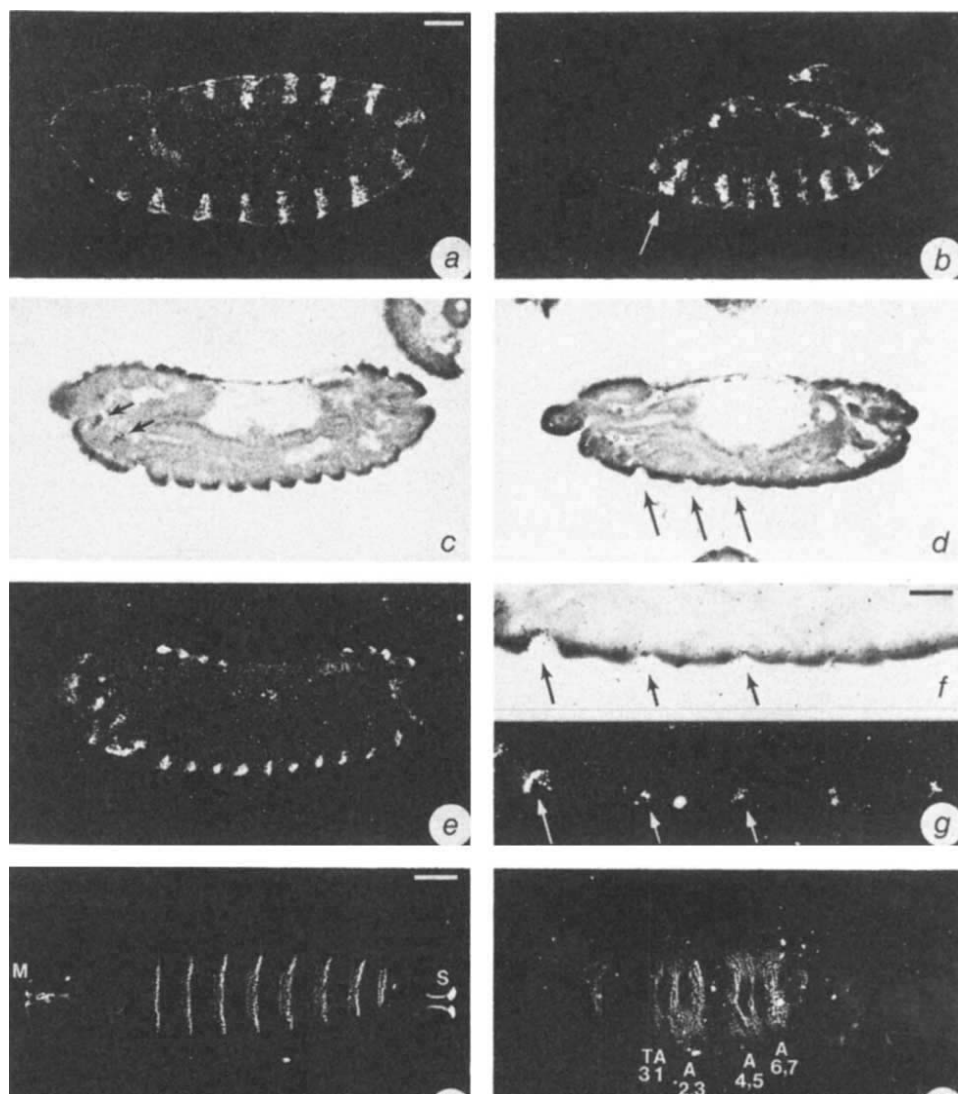
Fig. 3 The order of appearance of the even-numbered bands of *engrailed* hybridization was determined in 247 section edges from 132 sections of 105 different embryos. Section edges chosen for examination were those with clear differences in band intensities. After a pairwise comparison of each of the even-numbered bands, only cases of clear intensity differences were tabulated. *b*, Total number of instances in which two bands with different intensities were compared; *a*, the percentage of such comparisons in which the bands designated along the horizontal axis were more intense than those indicated on the vertical axis. The boxed numbers indicate cases of posterior stripes appearing before anterior ones: 8 before 6 and 12 before 10. *c*, A summary of the types of patterns that were observed. Embryos of increasing age are portrayed as rectangles with their posterior pole to the right. The dotted areas indicate bands of hybridization.

four-segment reiterated pattern; and (3) when all of the even-numbered bands become prominent, they have autoradiographic signals that are briefly stronger than those of the intervening odd-numbered bands, suggesting a two-segment reiterated pattern. A detailed description of this progression follows.

The first of the prominent bands of hybridization was observed in early cycle 14 embryos whose nuclei had just initiated the elongation process (Fig. 1*c, d*). A stripe of grains with a width of one to two nuclei was oriented perpendicular to the long axis of the embryo, about one-third from the anterior end. Grain density was initially greater over the nuclei than over the adjacent cytoplasm (Fig. 1*c, d*), but during the 14th cycle, the relative proportion of cytoplasmic grains increased (Fig. 1*e, f*). With the commencement of gastrulation near the end of cycle 14, the stripe of hybridization was at the posterior edge of the cephalic furrow, and corresponded to the second of the 14 stripes from the anterior end (Fig. 2*f*). The posterior edge of the cephalic furrow is the position of the mandibular or maxillary head segment primordium^{9,19}.

The fourth and eighth bands, weak concentrations of grains at the beginning of cycle 14 (Fig. 1*c, d*), were (along with band 2) consistently the most intense as gastrulation began (Fig. 2*a-d*). The relative strengths of the autoradiographic signals over the fourth and eighth stripes were often unequal (82 cases). In such sections in which the ventral and dorsal surfaces could be unambiguously identified, band 4 was more intense than band 8 dorsally (19/19 cases), and band 8 was more intense ventrally (13/16 cases) (Fig. 2*a*). Thus, ventrally, band 8 appeared before 4, and dorsally, band 4 appeared before 8. Band 12 was also observed as a prominent early autoradiographic signal (Fig. 2*c, d*), although with less regularity.

Fig. 4 Micrographs of *engrailed* transcripts in *engrailed* mutants. **a**, Dark-field micrograph of a section of a gastrulating wild-type embryo. The grains of adjacent stripes are approximately equal in density (also see Fig. 2f). **b**, Dark-field micrograph of a section of an *en^{C2}/en^{SF31}* mutant embryo during germ-band elongation (*en^{SF31}* deletes the entire *engrailed* locus). Signal alternates in strength in adjacent segments and the phasing is the same as observed in younger wild-type embryos: prominent stripe 2 is located in the posterior part of the cephalic furrow (arrow). **c**, **e**, Bright-field and dark-field images of a wild type embryo following germ-band shortening. The signal strengths are approximately equal over adjacent segments. Note the grains over the anterior portion of the ventral nerve chord (arrows in **c**). Serial sections of the embryo revealed labelling over the other ventral nerve chord ganglia. Such expression is consistent with earlier reports that *engrailed* function is essential in the adult nervous system²⁸. **d**, Bright-field micrographs of a section of an *en^{LA4}/en^{LA7}* mutant embryo. Pairs of segments are fused such that deep grooves are only observed every second segment (arrows). **f**, **g**, Higher magnification views of the periphery of the mutant in **d**. Grains are over every second segment over the deep grooves (arrows). Few or no grains are over the sites of the segment fusions. **h**, Dark-field micrograph of the ventral cuticle of an 18-h wild-type embryo. Note the mouthparts (M), eight abdominal denticle bands, and posterior spiracles (S). **i**, Dark-field micrograph of the ventral cuticle of an *en^{LA4}/en^{SF31}* embryo. Fused abdominal denticle belts are indicated. Scale bars, 50 μ m in **a-e**; 25 μ m in **f**, **g**; and 100 μ m in **h**, **i**.



In 18 sections with intense labelling in bands 4 and 8 or 8 and 12, the strengths of the bands in the three-band interval between them varied such that the intervening odd-numbered bands were more intense than the even-numbered ones (Fig. 2b).

To document the order of appearance of these even-numbered bands, sections whose bands varied in intensity were examined. In 132 such embryo sections, pairs of even-numbered bands were compared and instances of clear differences in intensity were scored (Fig. 3). Band 2 was almost always more intense, and with only two exceptions, the more anterior bands were more intense than each of the posterior bands. As exceptions, the 8th was usually more intense than the 6th and the 12th more than the 10th (Fig. 3a, b). A pictorial representation of embryos with the implied four-segment periodic patterns is presented in order of increasing complexity and maturity (Fig. 3c, see also Figs 1d, 2b-d).

Even-numbered autoradiographic stripes increased in intensity throughout the 14th cycle. Between these were stripes with less intense signals. The resulting pattern, most prominent at the beginning of gastrulation, had a two-segment periodicity with alternating strong and weak autoradiographic signals over adjacent segments (Fig. 2e). The autoradiographic grains of the more intense even-numbered bands were dense over both the nuclear and cytoplasmic regions; in contrast the less intense odd-numbered bands were predominantly labelled over the nuclei.

The alternating pattern leads within minutes to an arrangement of stripes with equal intensity. The equalization process

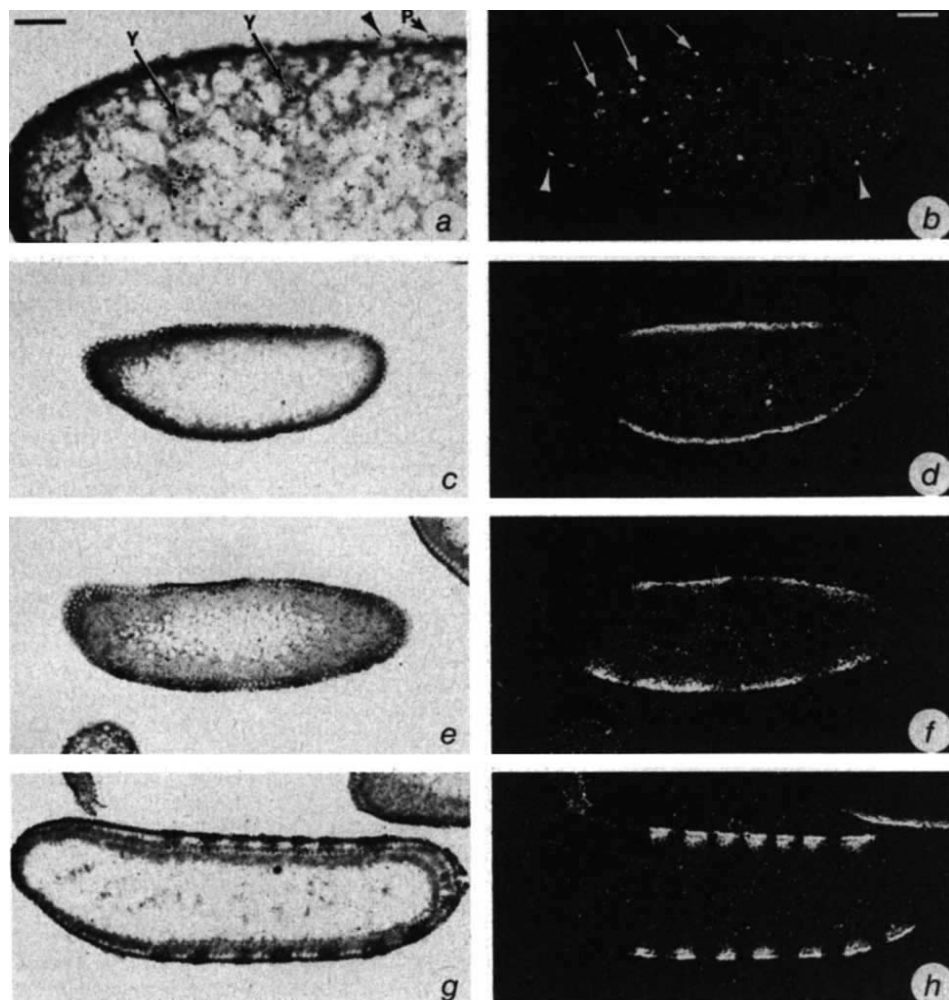
occurred with an anterior to posterior and a ventral to dorsal progression. As gastrulation proceeds, the ventral germ band extends around the posterior end onto the dorsal surface. In sections of germ band elongation embryos, additional bands of hybridization in the head and posterior regions were observed as reported previously². By this stage, all of the stripes were two to three cells wide and were equally intense. Consistent with previous descriptions, cells expressing the *engrailed* locus were observed in the progenitors of the mesoderm and nervous system throughout the germ-band elongation stages. However, the improved preservation of sections and longer exposures revealed, in addition, expression in each segmental ganglion of the embryonic nervous system during later stages (Fig. 4c).

Expression of *fushi tarazu* in early embryos

Gene expression in a pattern with reiterated periodic bands was first described for the *fushi tarazu* locus⁸. To compare the processes that generate the patterns of *engrailed* and *ftz* expression, embryo sections were hybridized with a *ftz* probe in parallel to the experiments described above for the *engrailed* gene. Our results largely confirm the work of Hafen *et al.*⁸, but reveal two additional intermediate stages in the temporal and spatial pattern of *ftz* expression—generally dispersed transcription at nuclear cycle 10 and a four-band stage early in cycle 14.

Localization of *ftz* RNA was first detected in syncytial blastoderm embryos during their 10th nuclear division cycle. Autoradiographic grains were localized over both yolk and

Fig. 5 Patterns of *fushi tarazu* transcripts in syncytial and cellular blastoderm stage embryos. *a, b*, *ftz* transcripts were detected over peripheral (P) and yolk nuclei (Y) at cycle 10 (arrows in *a* and *b*). The bright-field micrograph is a high-magnification image of the section shown in dark-field in *b*. Arrows point to corresponding nuclei. Arrowhead in *a* points to an unlabelled nucleus. Arrowheads in *b* point to examples of labelled peripheral nuclei. *c, d*, Embryo at cycle 12. During cycles 11–13, *ftz* transcript is evenly distributed over the middle portion at the edge of the embryo as previously reported⁸. *e, f*, Cycle-14 embryo; *ftz* transcripts are in four broad bands. *g, h*, Cycle-14 embryo with membrane furrows extended just beyond the nuclei; *ftz* transcripts are in seven bands. The probe was pDmV61H3.5 (a genomic 3.5 kilobase pair *Hind*III subclone in pUC8²⁹), nick-translated with tritiated nucleotides. Scale bar, 25 μ m in *a*, and 50 μ m in *b–h*.



peripheral nuclei (Fig. 5*a, b*). Labelled nuclei were along the entire egg length, although not all peripheral nuclei were labelled. By the 12th cycle, *ftz* RNA was uniformly distributed in the cortical region of the embryo, from approximately 15% to 65% egg length (Fig. 5*c, d*). No further regional localization of the hybridizing RNA was observed until the 14th cycle, when autoradiographic grains were in discrete regions that approximate four broad bands (Fig. 5*e, f*). The most anterior band was not as wide as the others, but as the nuclei elongate and before the membrane furrows penetrate into the embryo, the three posterior broad stripes subdivided and yielded a pattern of seven discrete regions of hybridization (Fig. 5*g, h*). At both the four- and seven-band stages, the discrete gaps that define the bands appeared in an anterior-to-posterior and ventral-to-dorsal progression. There was no apparent difference in density of autoradiographic grains over the different stripes. In stage 14 embryos *ftz* transcripts were more numerous than *engrailed* as estimated from grain densities after *in situ* hybridization.

The existence of the transient four-band stage was documented by tabulating the order of appearance of gaps between hybridization stripes. With only three exceptions, the gaps defined the more anterior stripes first. These exceptions were: the gap between stripes 3 and 4 formed before the stripe 2–3 separation, and the gap between stripes 5 and 6 formed before the stripe 2–3 and stripe 4–5 separations (Fig. 6). Figure 6*c* summarizes the implied progression of *ftz* expression from uniform to banded distribution.

To establish the relationship between the stripes of *ftz* and *engrailed* hybridization and the segment primordia in the blastoderm, the repeat lengths of the various periodic patterns were

compared. In sections of early stage 14 embryos in which the gaps between bands of *ftz* transcript were developing, the distance between gaps 1–2 and 3–4, and between gaps 3–4 and 5–6 was $88 \pm 11 \mu\text{m}$, corresponding to 13.8 ± 1.8 adjacent nuclei (56 measurements were taken). In sections of early stage 14 embryos hybridized with the *engrailed* probe, the distance between prominent bands 4 and 8, and between 8 and 12 was $92 \pm 11 \mu\text{m}$, corresponding to 14.4 ± 1.7 adjacent nuclei (46 measurements). Thus, the intervals of *engrailed* expression defined by bands 4, 8 and 12 have the same periodicity as the initial gaps of *ftz* expression. Given the estimated segment width of 3–4 cells⁹, this periodicity corresponds to four segments. As described previously^{2,3,8}, later cycle 14 embryos with 14 stripes of *engrailed* and 7 stripes of *ftz* expression have periodicities that correspond to one and two segments, respectively.

Throughout the syncytial blastoderm and cellular blastoderm stages, the autoradiographic grains of *ftz* hybridization were most concentrated over the periplasmic region between the nuclei and the embryo cortex. This sequestration of *ftz* RNA contrasts with the more generally dispersed *engrailed* RNA and may suggest novel mechanisms for the regulation of RNA storage and utilization.

Transcript localization in *engrailed* embryos

All developmental stages^{18,20–22} are affected by *engrailed* mutations. The *engrailed* cells in mosaic adults develop abnormally in all posterior compartments and generate similar types of defects in each segment^{20–24}. Yet mutant embryos have fused pairs of segments, and very occasionally, fusions of groups of four segments (Fig. 4*h, i*)²². A gene such as *engrailed* that is

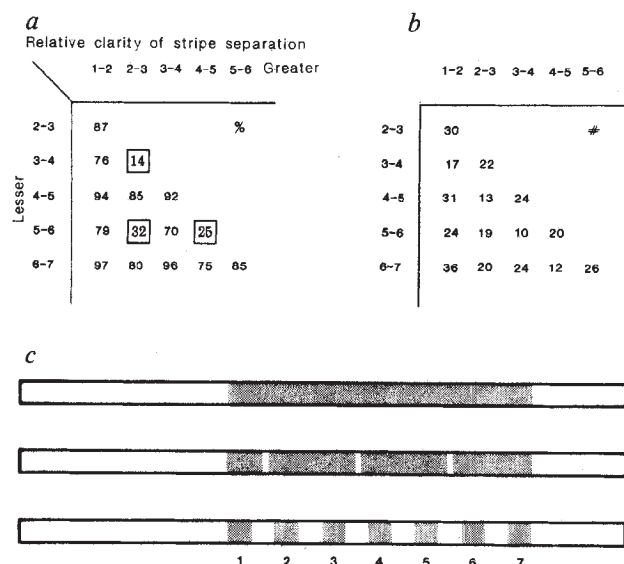


Fig. 6 The order of appearance of gaps in *ftz* expression along the embryo periphery was determined in 68 section edges from 34 sections. Sections chosen for analysis had gaps of varying clarity. For each comparison of two gaps, part *b* indicates the total number of cases with differences in the clarity of the gaps and part *a* indicates the percentages of such comparisons in which the gaps designated along the horizontal axis were more pronounced than the gaps on the vertical axis. The boxed figures indicate cases of posterior gaps becoming pronounced before anterior ones. *c*, A summary of the types of patterns observed. Embryos of increasing age are portrayed as rectangles with their posterior pole to the right. The dotted areas indicate regions of hybridization.

Discussion

The surface of the *Drosophila* cellular blastoderm is uniform in appearance, but its constituent cells are not equivalent. These studies and previous work⁸ have shown that during the 14th nuclear division cycle, the blastoderm cells differentiate to express selected genes in remarkably precise patterns. Maturation of the expression patterns of the *engrailed* and *ftz* genes have provided two views of this differentiation process, a process that progressively subdivides the embryo into biochemically and spatially distinctive portions.

Apparent differences between the development of the *engrailed* and *ftz* patterns of expression include earlier maturation of the *ftz* pattern and contrasting modes for generating bands of expression: new bands of *engrailed* expressing nuclei seem to be added while *ftz*-expressing nuclei are suppressed in a patterned arrangement. However, direct comparisons between the patterns of *engrailed* and *ftz* RNA are made difficult by the lower abundance of the *engrailed* transcripts. The presence of *engrailed* transcripts in cleavage and syncytial blastoderm stages has been demonstrated by Northern analysis¹⁸, but the low-level expression, which may be generally dispersed or distributed in a portion of the embryo, was not detectable with our *in situ* hybridization methods. The appearance of the *engrailed* patterns in *in situ* autoradiograms may be delayed until transcript levels are sufficient to be detected by our assay. Therefore, the great quantitative difference between levels of *ftz* and *engrailed* transcripts may mask fundamental similarities.

For both *engrailed* and *ftz*, the mature pattern of equally sized and equally spaced stripes of expressing cells emerges from patterns with longer repeat lengths that define fewer units. Maturation proceeds with anterior to posterior and ventral to dorsal polarities. For *ftz*, expression progressively narrows to four broad bands with a four-segment repeat length before dividing into seven narrower ones with a two-segment-repeat length. For *engrailed*, bands one to two cells wide define long intervals that shorten in a stepwise manner—first two bands (2 and 8 ventrally), then a four-segment periodicity, then two, and then in one-segment intervals.

That reiterated patterns of expression progress from longer to shorter repeat lengths suggests that the process of segmentation involves progressive partitioning of large units into smaller ones. The existence of discrete intermediate stages of segmentation has been proposed previously on theoretical grounds²⁶, and is strongly supported by the pattern of *engrailed* expression in *engrailed* mutants. Mutant late gastrulas had *engrailed* transcripts in patterns characteristic of younger embryos—in intervals greater than a single segment—indicating that the segmental fusions of *engrailed* mutants may be relics of earlier developmental units. A similar conclusion was drawn earlier from the phenotypes of the 'pair-rule' segmentation mutants that affect homologous portions of every second embryonic segment. It was suggested that such phenotypes are indicative of the transient existence of double-segment-sized units or fields²⁵.

The sequence in which the spacing of segments in the embryo is defined eliminates several possible mechanisms of segmentation. It indicates that the precisely spaced rows of cells in each segment are not designated in a simple sequential fashion at some predefined interval along the embryo's longitudinal axis. Nor do the many elements of the pattern appear simultaneously in their final form. Rather the mature pattern seems to involve proportioning of the embryo into successively smaller units. Such proportioning may provide the basis for coding position along the embryo's longitudinal axis. The overlapping activities of genes expressed with periodicities of one segment, two segments and larger intervals, together could direct in a combinatorial way the development of the blastoderm cells.

This work was supported by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research to M.P.W. and a NIH grant to T.K. We thank M. Scott for providing the *fushi tarazu* genomic clone, T. Karr and M. Simon for technical advice, and S. Poole and G. Martin for helpful comments on the manuscript.

expressed in every segment might be expected to affect all segments similarly, whereas pairwise segment fusions might be expected in mutants of genes required in every second segment²⁵. Thus the *engrailed* embryonic phenotype was surprising. We now demonstrate a direct correlation between the fusion phenotype of *engrailed* mutant embryos and the expression of the *engrailed* gene in periodic, two-segment-wide intervals.

Sections of *engrailed* mutant embryos were hybridized with the *engrailed* complementary DNA probe. Mutants examined were *en*^{LA4}/*en*^{LA7} heterozygotes (both apparent point mutations), and *en*^{C2} (an inversion breakpoint within the *engrailed* locus approximately 15 kilobase pairs upstream of the transcription unit (B. Drees and T.K., unpublished)) hemizygotes²². Mutants were identified in sections by their abnormal morphology and abnormal distribution of *engrailed* transcripts. Whereas in wild-type embryos the intensity of hybridization to the posterior compartment stripes alternated for a brief period in the cellular blastoderm and early gastrula and was uniform thereafter (Figs 2f, 4a, c, e), in most mutant embryos the alternating intensity persisted through germ-band elongation and germ-band shortening (Fig. 4b, d, f, g). The identities of the more prominent stripes in the mutant embryos corresponded with the location of the more intense stripes of hybridization of younger wild-type embryos; for example, the stripe at the posterior edge of the cephalic furrow was prominent in both (arrow in Fig. 4b). Comparison of the phasing of the alternating transcriptional pattern with the fused segment morphology of the mutants indicated that the weaker stripes of the alternating pattern coincide with the positions of the segment fusions (Fig. 4d, f, g).

Thus, apparent point mutations or breakpoint mutations upstream of the transcription unit both affect the spatial and temporal pattern of *engrailed* expression. This demonstrates that DNA 5' to the transcription unit is involved in transcriptional regulation and that the *engrailed* function may modulate its own expression, either by directly affecting transcription or through interactions with other regulatory factors.

Received 4 September; accepted 2 October 1985.

1. Garcia-Bellido, A. *Ciba Fdn Symp.* **29**, 161-182 (1975).
2. Kornberg, T., Siden, I., O'Farrell, P. & Simon, M. *Cell* **40**, 45-53 (1985).
3. Fjose, A., McGinnis, W. & Gehring, W. *Nature* **313**, 284-289 (1985).
4. Wieschaus, E. & Gehring, W. *Dev Biol.* **50**, 249-263 (1976).
5. Lawrence, P. & Morata, G. *Dev Biol.* **56**, 40-51 (1977).
6. Steiner, E. *Wilhelm Roux Arch. EntwMech.* **180**, 9-30 (1976).
7. Wakimoto, B. & Kaufman, T. *Dev Biol.* **81**, 51-64 (1981).
8. Hafen, E., Kuroiwa, A. & Gehring, W. *Cell* **37**, 833-841 (1984).
9. Lohs-Schardin, M., Cremer, C. & Nusslein-Volhard, C. *Dev Biol.* **73**, 239-255 (1979).
10. Rabinowitz, M. *J. Morph.* **69**, 1-49 (1941).
11. Sonnenblick, B. in *Biology of Drosophila* (ed. Demerec, M.) 62-167 (Wiley, New York, 1950).
12. Zolotar, M. & Erk, I. *J. Microsc. Biol. Cell* **25**, 97-106 (1976).
13. Turner, F. & Mahowald, A. *Dev Biol.* **50**, 95-108 (1976).
14. Foe, V. & Alberts, B. *J. Cell Sci.* **61**, 31-70 (1983).
15. Fullilove, S. & Jacobson, A. *Dev Biol.* **26**, 560-577 (1971).
16. Zolotar, M. *Dev Biol.* **49**, 425-427 (1976).
17. Lamb, M. & Laird, C. *Dev Biol.* **52**, 31-42 (1976).
18. Karr, T., Ali, Z., Drees, B. & Kornberg, T. *Cell* (in the press).
19. Underwood, E., Turner, F. & Mahowald, A. *Dev Biol.* **74**, 286-301 (1980).
20. Garcia-Bellido, A. & Santamaria, P. *Genetics* **72**, 87-104 (1972).
21. Lawrence, P. & Morata, G. *Dev Biol.* **50**, 321-337 (1976).
22. Kornberg, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1095-1099 (1981).
23. Kornberg, T. *Dev Biol.* **86**, 363-372 (1981).
24. Lawrence, P. & Struhl, G. *EMBO J.* **1**, 827-833 (1982).
25. Nusslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
26. Kauffman, S., Shymko, R. & Trabert, K. *Science* **199**, 259-270 (1978).
27. Mitchison, T. & Sedat, J. *Dev Biol.* **99**, 261-264 (1983).
28. Lawrence, P. & Johnston, P. *EMBO J.* **3**, 2839-2844 (1984).
29. Laughon, A. & Scott, M. *Nature* **310**, 24-31 (1984).

Transcription pattern of the *Drosophila* segmentation gene *hairy*

P. W. Ingham, K. R. Howard & D. Ish-Horowicz

Developmental Genetics Laboratory, Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, UK*

Segmentation of the Drosophila embryo requires expression of the pair-rule genes, mutations of which cause reiterated deletions in alternate segments along the antero-posterior body axis. We find that transcripts of one such gene, hairy, accumulate in eight distinct regions of the early embryo. This pattern of expression is compared with that of another pair-rule gene, fushi tarazu, and its dependence on maternally expressed genes is described.

A COMMON theme in the organization of many metazoa is the subdivision of the body into a series of homologous segments¹. Different structures are then elaborated in the various segments to generate complex forms. A vivid example of such segmental or metamer organization is provided by the insects. Although segmentation of the insect body has been studied for many years²⁻⁴, the prospect of an understanding at the molecular level has only recently come into view with the identification, in *Drosophila melanogaster*, of most of the genes required for the process⁵⁻⁹.

Metameric organization is established during early *Drosophila* embryogenesis¹⁰⁻¹² as the incipient cells of the blastoderm embryo respond to information generated epigenetically in the maternally derived environment of the egg^{13,14}. This response depends on the expression of the zygotic genome and, in particular of genes defined by a class of zygotically acting mutations⁵⁻⁸ which disrupt the generation of the normal segmental pattern. These fall into three distinct phenotypic classes: (1) the gap mutations, which cause the deletion of contiguous groups of segments, (2) the pair-rule mutations, which result in deletions of homologous pattern elements in alternate segments, and (3) the segment polarity mutations, which affect the patterning of every segment.

The discovery of the pair-rule class provided the first evidence of a double-segment order during the formation of the metamer pattern⁵. This order is reflected at the molecular level by the pattern of expression¹⁵ of the pair-rule gene *fushi tarazu* (*ftz*)^{16,17}. At the blastoderm stage, *ftz* transcripts accumulate in seven discrete regions which repeat along the antero-posterior axis of the embryo with double-segment periodicity¹⁵. Such a precise pattern of transcript distribution implies that the *ftz*⁺ function is specifically required for the establishment of particular metamer units—those which are deleted in *ftz*⁻ mutants. Mutants of a second pair-rule gene, *hairy* (*h*)^{5,18}, cause a pattern of deletions which is approximately complementary to that typical of *ftz*, suggesting a corresponding complementarity of spatial expression and function.

We have analysed the spatial distribution of *h* RNA in early embryos, and compare it with the previously described *ftz* transcription pattern. We also show that establishment of the *h* transcription pattern depends on maternal determinants of both antero-posterior and dorso-ventral polarity.

Effect of *hairy* on embryonic segmentation

The metamer organization of the *Drosophila* body is directly reflected in the cuticle of the first-instar larva. This is especially clearly illustrated by the pattern on its ventral surface (Fig. 1 and ref. 19). Each thoracic and abdominal segment secretes a characteristic belt of cuticular processes or denticles. The metamer origin of the mouthparts is obscured by head involution, but cuticular structures characteristic of each segment can be easily identified.

Larvae homozygous for *h* null mutations exhibit a range of segmentation defects¹⁸ representing variations on a common theme, the pair-rule phenotype. The most regular *h* phenotype (Fig. 1c) consists of the deletion of the anterior region of each even-numbered abdominal segment and the posterior region of each odd-numbered abdominal segment from the larval cuticular pattern. This is almost the precise complement of the phenotype characteristic of amorphic alleles of the pair-rule gene *ftz* (Fig. 1d). In both cases, the same register of deletions extends into the thoracic and head segments (Fig. 1a). In *h* mutants, structures which derive from the mandibular (G1) and labial (G3) segments (the ventral arms of the cephalopharyngeal skeleton and H-piece respectively)²⁰ are deleted. In addition the medial tooth, a derivative of the labrum, is disrupted. This contrasts with the phenotype of strong *ftz* alleles in which only derivatives of the post-oral maxillary and labial segments are affected.

This regular pair-rule phenotype is typical only of a proportion of animals homozygous for any given amorphic *h* allele¹⁸. Commonly, more extreme defects are seen. The entire third thoracic segment is often deleted such that the second thoracic and first abdominal dentical bands fuse; similarly the remaining three posterior abdominal bands frequently fuse to form a continuous 'lawn' of denticles exhibiting polarity perturbation^{18,21}.

These lesions in the larval cuticle are manifestations of defects occurring much earlier during embryogenesis. Absence of *h*⁺

* New address from December 1985: ICRF Developmental Biology Unit, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

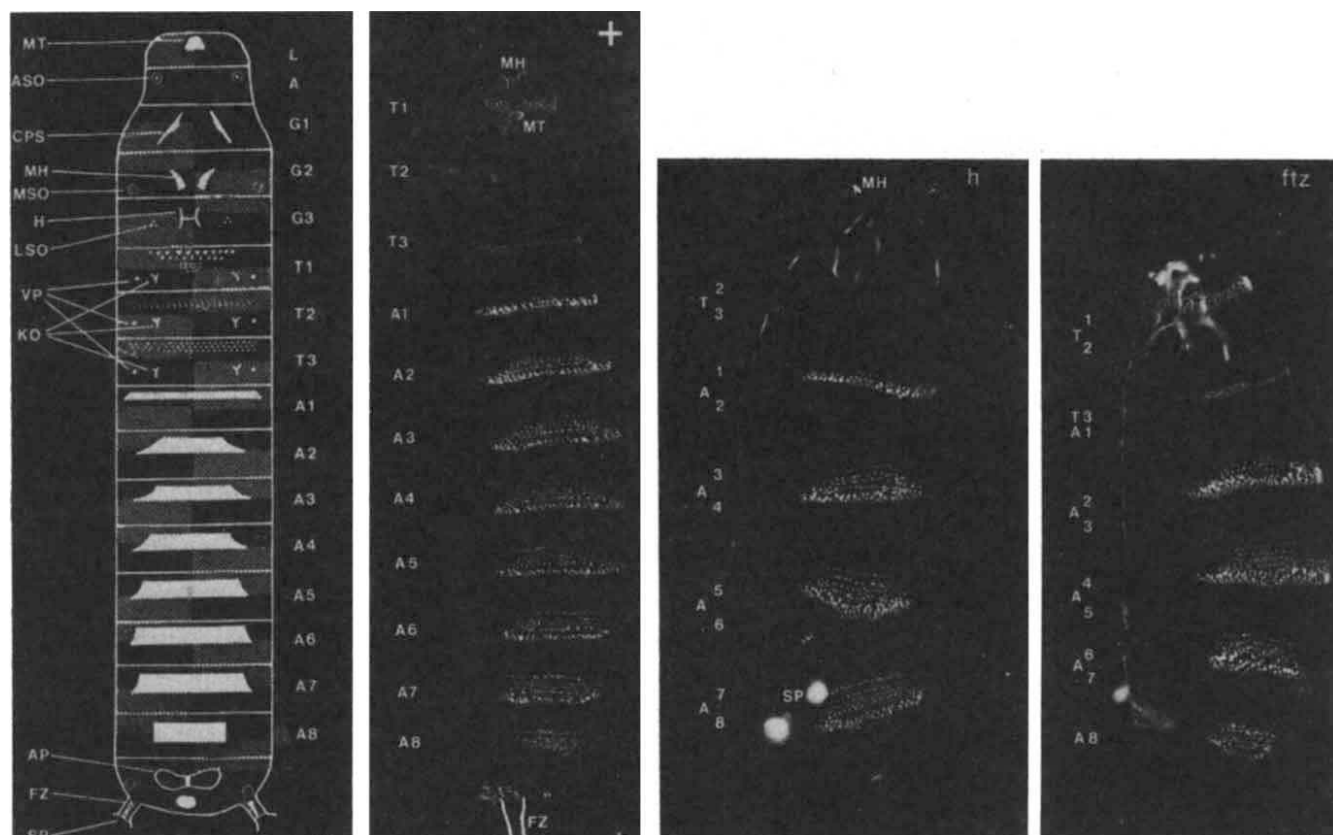


Fig. 1 *a*, Schematic representation of a first-instar *Drosophila* larva. The cephalon and mouthpart segments are drawn as they would appear in the absence of head involution. Each thoracic and abdominal segment is characterized by a ventral denticle band, represented by white bands. The thoracic segments each bear a pair of ventral pits (VP) and Keilin's organs (KO) which probably mark the position⁴⁰ of the antero-posterior compartment⁴¹ boundary. The mouthpart segments each differentiate specific structures: LSO, labial sense organs; H, H-piece; MSO, maxillary sense organs; MH, mouth hooks; CPS, cephalopharyngeal skeleton. Anterior to these are the antennal sense organs (ASO) and the medial tooth (MT) of the labrum (see ref. 20). The hatched regions represent the consensus deletions typical of either *h* (left) or *ftz* (right) mutations (see *c* and *d*). *b*, Ventral cuticle of a first-instar larva viewed under dark-field illumination, showing the thoracic (T1-T3) and abdominal (A1-A8) denticle bands. *c*, Larvae homozygous for a deficiency or null mutations of *h* exhibit a reduction in the number of ventral denticle bands. Specifically, most of the denticle bands of each even-numbered abdominal segment (A2, A4, A6, A8) are deleted. In addition the naked posterior parts of each odd-numbered segment (A1, A3, A5, A7) are deleted. This frame of deletion extends into the thoracic and gnathal segments; in addition the medial tooth, a derivative of the labrum, is disrupted. The precise limits of the deleted regions are indicated in the left-hand half of the schematic larva shown in *a*. These have been deduced from the study of partial deletions in weaker genotypes (P.W.I. *et al.*, in preparation) and, in the case of the A1/A2 deletion, by using the *Ultrathorax* (*Ubx*) mutation to mark anterior A1 derivatives²³. The deleted region in larvae exhibiting this regular pair-rule phenotype varies in size but is approximately the same width as a segment. It includes the segment boundary, its limits being anterior to the presumed position⁴⁰ of the anterior/posterior compartment boundary in each segment. *d*, Larvae carrying null mutations of *ftz* (*ftz*^{W20/ftz^{9H34})^{7,16} exhibit a phenotype which is the reciprocal of that typical of *h* mutants. Thus the anterior of each odd-numbered and the posterior of each even-numbered abdominal segment is deleted. This deletion frame similarly extends into the thoracic and gnathal segments and is shown schematically in the right-hand half of *a*. Our assessment of the *ftz* phenotype differs in several respects from previously published descriptions^{16,17}. For instance we find the H piece, a derivative of the labial segment, to be present, in contrast to the assertion of Wakimoto *et al.*¹⁶ (but see their Fig. 6d which clearly shows an H piece). We also consider the boundaries of the deletion frame to lie, like those of *h*, anterior to the anterior/posterior compartment boundary. This is supported specifically by three observations: (1) The secondary patch of denticles in T1 which lies posterior to the major denticle band but anterior to the Keilins organs (which mark the anterior/posterior boundary)⁴⁰ is partially deleted in *ftz* larvae. (2) The Keilin's organs posterior to the T3 denticle band are usually deleted. (3) Part of the salivary gland which derives from the anterior compartment of G3 is retained in *ftz* mutants (P.W.I., K.R.H., M. E. Akam and A. Martinez-Arias, in preparation).}

has a clearly visible effect within 1 h of the cellular blastoderm stage when the germ band fails to extend fully (data not shown). By the time the parasegmental grooves²² are visible (Fig. 2*a*) the metamer organization is obviously disturbed, the germ band being divided into, at most, only half the normal number of parasegmental units (Fig. 2*b*). At the same time, discrete regions of cell necrosis become evident along the length of the germ band (Fig. 2*c*).

hairy RNA during early embryogenesis

Several different transcripts derive from the *h* gene (ref. 23 and K.R.H., in preparation). During embryogenesis two species can be detected by Northern blot hybridization: an RNA of about 2.1 kilobases (kb), denoted α_L , and, at lower levels, a slightly smaller (~2.0 kb) transcript, α_S . The expression of the *h* α transcripts during early embryogenesis has been analysed by *in situ* hybridization of a 1.7-kb genomic probe to wax-embedded sections. Hybridization to adjacent sections with *ftz* probe allows a direct comparison of the expression of both genes in the same embryos (see Fig. 3).

Both *h* and *ftz* transcripts are first detectable during the beginning of the interphase two cycles before cellularization of the blastoderm (stage 12)²⁴. Whilst *ftz* expression is restricted to a specific region, *h* transcripts are more or less uniformly distributed throughout the egg (Fig. 3*d-f*). Subsequently, two distinct regions of *h* transcript accumulation can be detected. Anteriorly, labelling is restricted to the dorsal half of the embryo in a region between 95% and 85% egg length, whilst a second domain of continuous labelling around the entire circumference of the embryo extends from approximately 75% to 20% egg length. Labelling in this latter domain rapidly becomes discontinuous (Fig. 3*g-i*); by mid-stage 14, prior to the completion of cellularization, *h* transcripts are localized in eight distinct regions along the antero-posterior axis of the embryo (Fig. 3*j-l*). The most anterior of these (designated region 0, see Fig. 4*a*) lies between 95% and 85% egg length, extending over 12–15 nuclei. The next most anterior region (region 1) extends around the entire circumference of the egg, being considerably broader ventrally (6–7 nuclei) than dorsally (3–4 nuclei). The remaining six regions (2–7) also extend around the whole circumference,

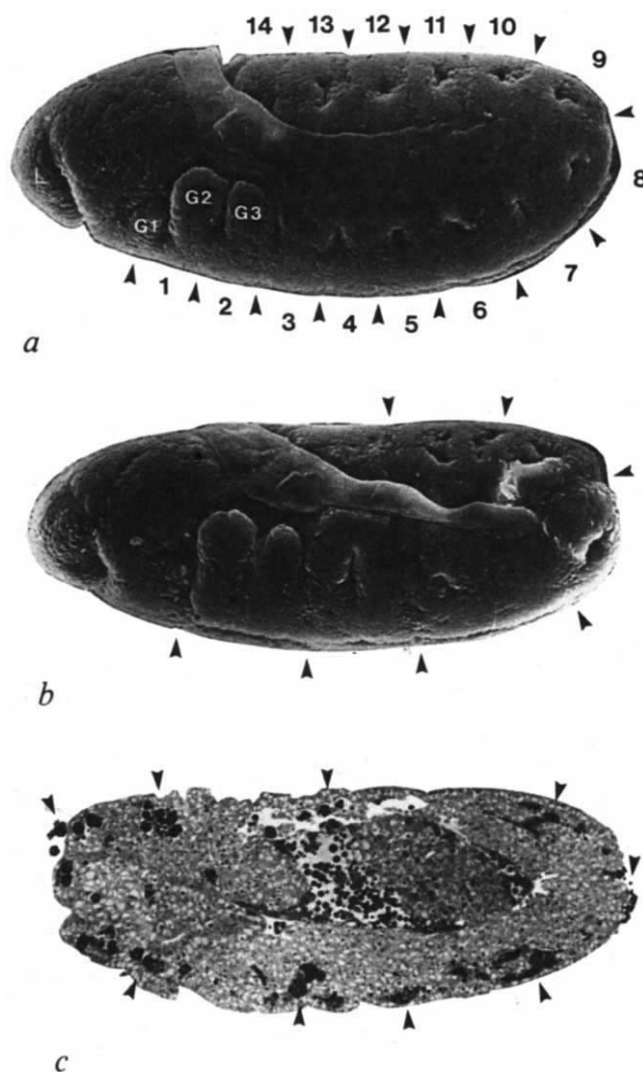


Fig. 2 In normal development, formation of the cellular blastoderm is followed almost immediately by the onset of gastrulation. Cells along the ventral midline invaginate to form the presumptive mesoderm and the ventral region of the embryo then begins to extend around the posterior pole. On completion of this process (after about 4.5 h of development at 25 °C) the embryo is folded back on itself with its posterior-most tip located dorsally just posterior to the initial position of the cephalic fold, the 'elongated germ band' stage. After a further 90 min, the first evidence of metamerism is visible⁴². The extended embryo becomes subdivided into 14 regions by superficial grooves in the ectoderm. These regions correspond to parasegments^{22,39}, the posterior compartment⁴¹ of one segment and the anterior compartment of the succeeding segment. **a**, A scanning electron micrograph of a wild-type embryo at the extended germ band stage (after about 6 h of development). Superficial grooves in the ectoderm (indicated by arrows) mark the boundaries between parasegments 1–14^{22,39}; the lobes correspond to the rudimentary gnathal appendages (G1–G3)⁴². The single large anterior-most lobe corresponds to the labrum (L). Within each of parasegments 4–13 is a deep invagination of cells which will form the tracheal system. These correspond to the position of the segment boundaries²². **b**, An embryo homozygous for the null allele¹⁸ *h*^{IL79K} at a similar developmental stage to that shown in **a**. This embryo is divided into only half the normal number of parasegments and possesses a correspondingly reduced number of tracheal pits. **c**, A lateral sagittal 2- μ m section through an embryo similar to that shown in **b**. Clusters of darkly stained dead cells are located in discrete regions along the length of the germ band (arrows). All specimens are oriented dorsal side up, anterior to the left.

and are of similar width (each 3–4 nuclei) and evenly spaced. Together, regions 1–7 extend from 75% to 20% egg length (Fig. 4d). At the same stage as the periodic pattern of *h* expression is fully resolved, transcripts from the pair-rule gene *ftz* are localized in seven stripes in the blastoderm (ref. 15 and

Fig. 3). These stripes are similar in width to those of *h* transcript at the equivalent stage.

Although some *h* transcript can be detected over the nuclei, much of the signal is located in the outer cytoplasm (Fig. 4c). This peripheral location of *h* transcripts in the cortical cytoplasm resembles that of *ftz* transcripts and contrasts with predominant localization of *Ubx* transcripts in the nucleus and inner cytoplasm²⁵.

Almost immediately after cellularization is complete, a transverse invagination of cells occurs one-third of the way from the anterior pole at 65–70% egg length¹⁹ (EL; the anterior pole is at 100% EL). This cephalic fold is the first morphological marker to appear along the antero-posterior axis of the previously uniform cellular blastoderm. The invaginating cells correspond to the primordia of the G1 and G2 segments^{19,26,27}.

At the onset of gastrulation, *h* regions 1 and 2 can be seen to flank the cephalic fold, region 1 being drawn into the anterior margin of the fold, whilst region 2 lies posterior to it (Fig. 4e). At the same stage the most anterior *ftz* region is located at the posterior of the fold (ref. 15 and Fig. 4f). As gastrulation proceeds, the periodically distributed *h* transcript rapidly decays (Fig. 3m–o). By the time the germ band has elongated, there is no evidence of the repeating pattern typical of the blastoderm. Transcript, however, is present in the hindgut and the foregut (Fig. 3p, q).

Domains of *h* and *ftz* RNA localization

To determine more precisely the relationship of the *h* and *ftz* patterns, probes for both transcripts were applied simultaneously to the same section so that the resultant signal represents the sum of the two individual signals (Fig. 5a, c). The two probes were also applied individually to the adjacent sections so that in any given embryo the signal from either *h* or *ftz* RNA alone could be distinguished (Fig. 5b, d). The mixed probes (Fig. 5a) clearly define the anterior dorsal patch of *h* RNA. Ventrally in the region between 75% and 15% EL, eight discrete regions of labelling can be resolved, each separated by a single unlabelled incipient cell. The eight labelled stripes are each about six or seven nuclei wide.

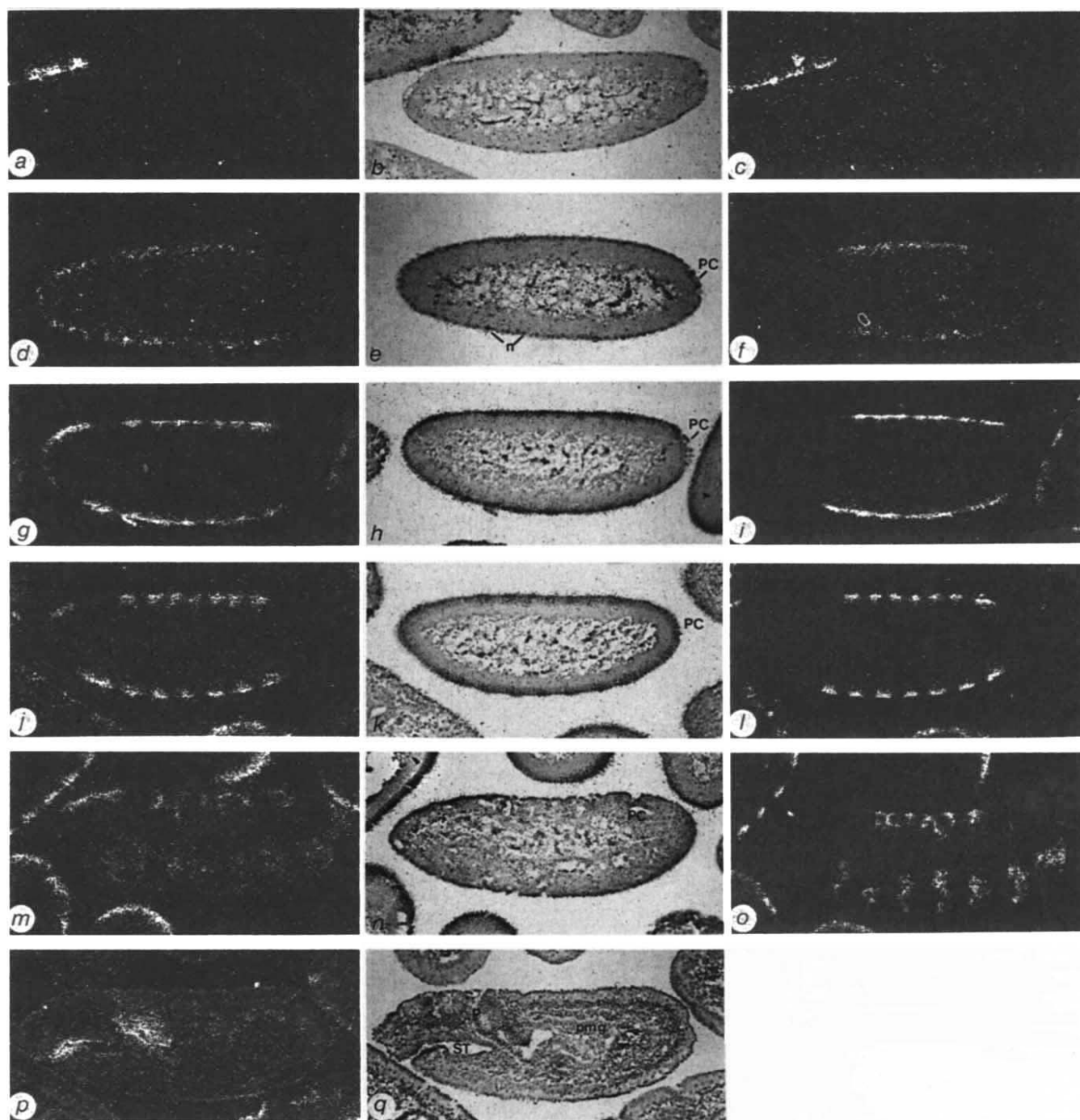
Since both *h* and *ftz* patterns consist of seven regularly spaced stripes, it is probable that the two terminal stripes in the pattern detected by the mixed probes consist of only one species of transcript. The anterior-most stripe in the mixed pattern represents *h* transcript, since *h* region 1 lies in the anterior part of the cephalic fold, whereas the anterior-most *ftz* stripe lies in the posterior part of the fold (see Fig. 4c, d). The posterior-most stripe must therefore be due to *ftz* transcript. Thus, the total domains defined by *h* and *ftz* transcripts are out of register by one metameric unit. In addition, the two patterns are out of phase with one another, though not to such a degree that they define complementary regions of the cellular blastoderm. The central six stripes must represent the addition of the two signals, with the anterior of each broad stripe being due to *ftz* signal and the posterior to *h* signal.

Maternal determinants

Polarity along the dorso-ventral and antero-posterior embryonic axes is established under the influence of maternally derived gene products¹³. The *h*-transcription pattern is asymmetrical along both these axes, implying that its generation requires such maternally acting genes. To test this possibility, *h* expression was analysed in embryos produced by two maternal-effect mutations, *dorsal* (*dl*) and *bicaudal* (*bic*).

Females homozygous for the *dl*¹ mutation produce embryos in which cellular fates are changed such that all regions of the embryo behave as though they were dorsal¹³. In the absence of the *dl*⁺ product, the *h* transcription pattern at stage 14 is altered, being approximately symmetrical along the dorso-ventral axis; this is particularly clearly illustrated at the anterior pole where transcript accumulates both dorsally and ventrally (Fig. 6b).

Homozygous mutant *bic* females produce embryos in which the anterior half is frequently replaced by a mirror-image



posterior^{13,28}. This change is invariably accompanied by a reduction in the total number of segments (see ref. 28 and Fig. 6c). Both these features are reflected in the pattern of *h* transcript distribution in stage 14 embryos derived from *bic* females. Region 0, which characterizes the anterior pole of the embryo, is frequently absent. Fewer regions of transcript accumulation are present (Fig. 6d), and their position differs from wild type.

An analogous effect on *ftz* expression has recently been reported²⁹.

Discussion

Transcription of *hairy* is first detectable during embryogenesis at the syncytial blastoderm stage. The transcripts, which are initially distributed throughout the embryo, become localized periodically along its antero-posterior axis at about the time the

body plan is thought to be established. This localization is presumably intrinsic to the process of segmentation because absence of *h*⁺ activity during embryogenesis causes a similarly periodic pattern of deletions along this axis. These deletions arise following a failure of the subdivision of the early embryo into the appropriate number of metameric units and are associated with localized cell death.

There are eight regions of *h* transcript accumulation in the cellularizing blastoderm, equivalent to the number of regions deleted from the larval cuticle. Superposition of the pattern of hybridization and the blastoderm fate map (Fig. 4d) suggests that these sites correspond broadly to the locations of the primordia of the missing structures. Similarly, *ftz* transcripts are localized in seven regions which correspond approximately to the primordia of structures deleted in *ftz* mutant larvae¹⁵.

Fig. 3 (Left) Comparison of *h* and *ftz* transcription during early embryogenesis. The central panels are bright-field images of sections through successively older embryos. In all cases, anterior is to the left. All sections are sagittal (dorsal side up) except for those in *d-f*. After nine rapid divisions in the yolk central region of the egg, most nuclei migrate to the peripheral cytoplasm to form the syncytial blastoderm^{24,43}. Here they divide a further four times, before cell membranes begin to grow inwards, isolating each nucleus and forming the cellular blastoderm (stage 14 of Foe and Alberts²⁴). Stages were determined by comparing the density of nuclei along the egg periphery. The pattern of *h* or *ftz* expression in each embryo is revealed by the dark-field images of adjacent sections hybridized with ³⁵S-labelled *h* probe (left) or ³H-labelled *ftz* probe (right) respectively. *a-c*, Stage 10 embryo; no expression of either *h* or *ftz* has been detected at this stage. Exposure, 30 days. *d-f*, Frontal sections of stage 13 embryo. Expression of *h* extends around most of the circumference of the embryo with the possible exception of the posterior polar region. In contrast, *ftz* expression is limited from the outset to a region from between 65% to 15% EL. In both cases the silver grains are most densely localized over the nuclei (n); the intensity of nuclear labelling is not uniform. PC, pole cells. Exposure, 30 days. The three embryos shown in the next nine panels are directly comparable being hybridized under precisely the same conditions. Exposure, 20 days. *g-i*, Early stage 14 embryo. Anteriorly, the *h* transcripts have become restricted to the dorsal surface. The remaining region is becoming discontinuous, with seven regions being apparent dorsally. At the same stage, the *ftz* transcripts are also starting to become localized. By mid-stage 14 (*j-l*), both *h* and *ftz* transcript patterns are well resolved into a series of stripes along the antero-posterior axis. This pattern begins to decay with the onset of gastrulation (*m-o*), the decay appearing more rapid in the case of *h* (compare the signal levels in *m* and *o* with the neighbouring stage 14 blastoderm in the top left corner). Note, however, that the anterior dorsal patch of *h* transcript remains prominent at this stage. By the completion of germ-band extension the periodic pattern of both transcripts has decayed. Expression of *h* is now detectable in the developing hindgut and foregut (the protodaeal (P) and stomodaeal (ST) invaginations respectively) but not in the posterior midgut (pmg), and persists in this region throughout the elongated germ-band stage. *p, q* Show such an embryo just prior to the onset of germ-band shortening. Exposure, 11 days.

Methods. Eggs were collected from rapidly laying fly populations on yeast/agar medium and incubated at 25 °C until the desired developmental stage had been reached. Prior to fixing, the chorion was removed by incubation in 14% sodium hypochlorite. Fixation was in 4% paraformaldehyde in phosphate-buffered saline. The embryos were first pre-fixed by the phase partition method⁴⁴, vitelline membranes were then removed by the method of Mitchison and Sedat⁴⁵. After rehydration, embryos were fixed for a further 15 min, dehydrated and embedded in paraffin wax. Sections (6 µm) were cut and mounted on glass microscope slides coated with poly-D-lysine. After removal of the wax with xylene, sections were rehydrated and prepared for hybridization essentially as described elsewhere^{46,47} except that the slides were additionally treated with 0.25% acetic anhydride in 0.1 M triethanolamine pH 8.0 for 10 min at room temperature before the final dehydration⁴⁸. DNA templates for transcription were subcloned into pSP64 or pSP65⁴⁹ and ³H- or ³⁵S-labelled RNA synthesized in a 20-µl reaction containing 40 mM Tris-HCl pH 8.0, 6.0 mM MgCl₂, 10 mM dithiothreitol, 4 mM spermidine, 100 µg ml⁻¹ bovine serum albumin, 0.1% Triton X-100, 1 U µl⁻¹ RNasin (Amersham), 100 µg ml⁻¹ linearized template DNA, 500 µM ATP and CTP with 100 µM ³H-GTP and NTP (NEN) or 500 µM ATP, GTP and CTP with 10 µM ³⁵S-UTP (NEN) and 0.2 U µl⁻¹ SP6 polymerase (Anglian Biotechnology) incubated at 30 °C for 4 h. These reaction conditions were found to maximize the yield of full-length transcripts (K.R.H., unpublished observations). The reaction was terminated by the addition of DNase I to 5 µg ml⁻¹ followed by incubation at 37 °C for 30 min. Probes were size reduced to an average of 50–100 bases using alkaline hydrolysis at pH 10.4 for 40 min at 60 °C⁵⁰. Final specific activities were 1.1 × 10⁹ d.p.m. µg⁻¹ for the ³H-labelled probes and 1.2 × 10⁹ d.p.m. µg⁻¹ for those labelled with ³⁵S-thiophosphate. The templates have been described previously^{23,51}. These probes were used at a final concentration of ~0.3 ng per µl per kb in 50% formamide, 10% dextran sulphate, 0.3 M NaCl, 10 mM Tris-Cl, 10 mM NaPO₄ pH 6.8, 5 mM EDTA, 1 × Denhardt's, 10 mM dithiothreitol, 1 mg ml⁻¹ transfer RNA⁵⁰. A 20-µl aliquot of this solution was dispensed on each slide and spread over the sections with a siliconized glass coverslip. Slides were sealed in plastic boxes saturated with wash buffer (50% formamide, 0.3 M NaCl, 10 mM Tris-HCl, 10 mM NaPO₄ pH 6.8, 5 mM EDTA, 1 × Denhardt's) and incubated overnight at 50 °C. Coverslips were removed by immersing slides in wash buffer plus 10 mM dithiothreitol (WDTT). Slides were washed in fresh WDTT for 4–5 h at 50 °C and treated with RNase A at 20 µg ml⁻¹ in 0.5 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA (NTE) for 30 min at 37 °C⁵⁰. Subsequently, they were washed in NTE for 1 h before a further final overnight incubation in WDTT at 50 °C. After dehydration of the sections, slides were coated with Kodak NTB2 emulsion, dried, stored at 4 °C and developed with Kodak D-19 developer (1:1 dilution) for 2 min at 20 °C after the stated exposure time. Controls using sense-strand hybridizations and treatment of antisense-strand hybridizations with RNase A in low salt (data not shown) indicate that the signals are due to specific hybridization.

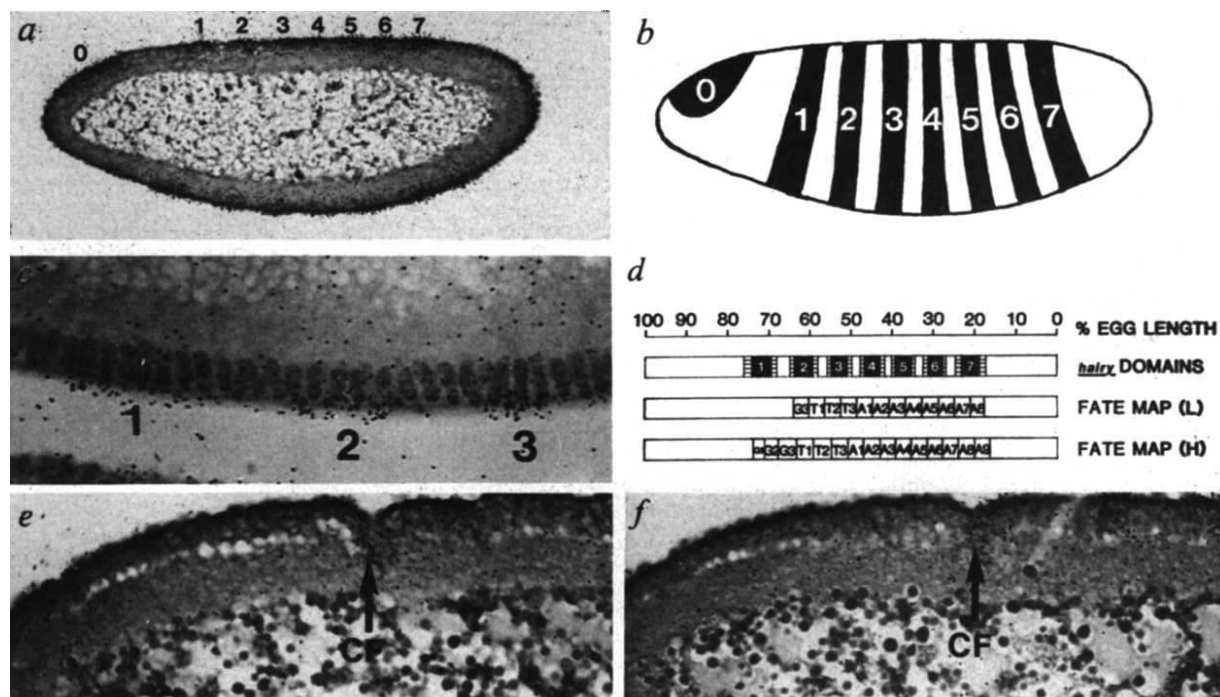


Fig. 4 *a*, Medial sagittal section of a mid-stage 14 blastoderm hybridized with ³⁵S-labelled *h* probe, revealing the eight regions (designated 0–7) of *h* transcript accumulation. *b*, A projection, onto the medial sagittal plane, showing the asymmetry of the pattern of *h* expression. This is based on several partial serial reconstructions. *c*, Detail of the ventral side of a late stage 14 blastoderm hybridized with ³H-labelled *h* probe. Note that domain 1 is broader than domains 2 or 3 which are both 3–4 cells wide. *d*, Projection of the regions of *h* RNA accumulation onto blastoderm fate maps (Lohs-Schardin *et al.*¹⁹ (L) and Hartenstein *et al.*²⁶ (H)). The filled regions show the average localization which involved medial frontal sections of six separate embryos (hatched regions indicate the s.d. of these measurements). Note the approximate correspondence with the primordia for regions deleted in the *h* mutant phenotype (compare with Fig. 1*a*). The two fate maps differ from one another in the exact location of the primordia. Although their estimates for the average size of the primordia (3.9% and 3.65%) are lower than ours (4.2 ± 0.16%), this is due mainly to the broadening of the *h* pattern seen anteriorly. Detailed comparisons with either fate map are not warranted since slightly different assumptions⁵² or level on the dorso-ventral axis¹⁹ were used to construct these maps. *e, f*, *In situ* hybridizations with *h* and *ftz* probes respectively to serial sections of the same early gastrula, illustrating the relationship of the regions of hybridization to the cephalic fold (CF).

One interpretation of these findings is that *h* or *ftz* are autonomously required for the specification of alternating sets of blastoderm cells; mutations in one or other gene would accordingly result in the death of these cells or their descendants. Since the *ftz* and *h* deletion patterns are approximately complementary, a *prima facie* expectation might be that the patterns of expression would show an equivalent relationship. Our results demonstrate, however, that even before cellularization is complete, some incipient cells in the metameric region of the blastoderm express neither gene at significant levels. Such cells could require the expression of one or more of the other pair-rule genes for their specification (see refs 15, 30).

We emphasize, however, that approximate correspondence between the sites of RNA localization and the primordia of structures deleted in the absence of *h* or *ftz* expression does not necessarily imply an autonomous requirement for either gene in these cells. An alternative possibility is that the pair-rule genes are a functionally heterogeneous class, the various genes acting in different ways to establish the metameric pattern.

Although the initiation and subsequent spatial distribution of *h* and *ftz* transcription follow a similar time course, they do exhibit some notable differences. Whereas *h* is initially expressed around most of the egg circumference, expression of *ftz* is restricted from the outset within the presumptive post-oral metameric region of the embryo. Moreover, the disposition of the *h* transcriptional regions appears less uniform than those of *ftz*; regions 0 and 1 are somewhat broader and region 4 slightly weaker than the remaining regions. These differences suggest that the two genes function at different levels.

The striped pattern of both *ftz* and *h* is unstable and has decayed within an hour of the cellular blastoderm stage, the decay of *h* appearing more rapid than that of *ftz*. This latter finding argues against a role for the pair-rule genes in maintaining the segmental pattern. Although it is conceivable that long-lasting protein products of these genes could effect such a function, it is more likely that they act as transient signals which regulate the expression of other genes. A primary candidate for such a gene is *engrailed* (*en*), which is first expressed in a periodic pattern in the cellular blastoderm^{31,32}. In contrast to both *h* and *ftz*, *en* expression continues throughout development in this spatially restricted pattern, reflecting its role in directing the development of the posterior compartment of each segment³³⁻³⁵.

The finding that *h* transcripts accumulate in the anterior region of the embryo throws some light on the question of head segmentation: region 0 includes the location of the labral primordium which is similarly restricted to the dorsal surface of the embryo. Although the lobed appearance of the labrum in the extended germ band (see Fig. 2a) is suggestive of an appendage and hence of its being segmental in origin, such an interpretation has previously been questioned on other morphological grounds³⁶. This would leave only one segment anterior to the mandibular (G1) segment, namely the antennal, the primordium of which maps between the *h* regions 0 and 1. Recently, however, evidence from fate mapping studies^{20,37} together with the analysis of homoeotically transformed larvae³⁸ has suggested the presence of at least two segments anterior to G1. In addition, *in situ* analysis of *engrailed* transcription clearly shows it to be expressed in two regions anterior to G1, one of which is located in the labrum^{3,39}. Taken together, these findings suggest a metameric origin for the labrum, with expression of *h* being required for its establishment.

The significance of the presence of *h* transcripts in the presumptive fore- and hindguts of the extended germ band is at present obscure. We note, however, that *en* is also expressed in these cells at the same stage^{31,32,39}.

The pattern of *h* RNA is not axially symmetrical, the stripes being broader ventrally than dorsally and inclined to the embryonic axis (see Fig. 4b). Presumably this reflects the true shape of the fate map at this stage, since such a pattern is seen well before the beginning of the movements of gastrulation. This contrasts with the assumption of a simple perpendicular

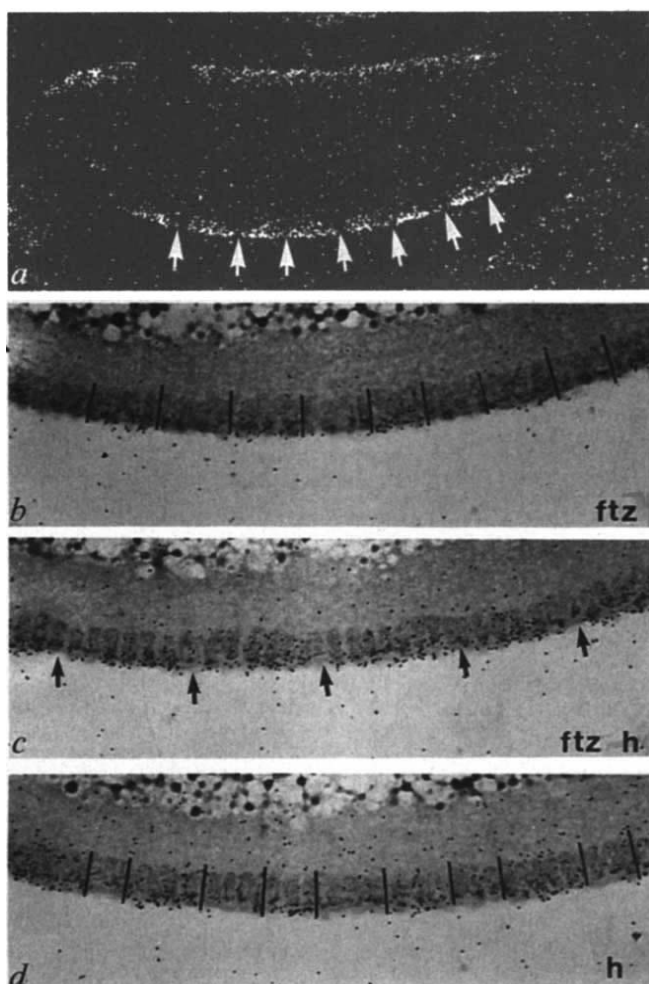


Fig. 5 The spatial relationship of *h* and *ftz* domains. The specific activities of the *h* and *ftz* probes were controlled such that each probe produced a similar signal intensity after a given exposure. Mixtures of the two probes, in equal concentrations, were applied to various sections of wild-type embryos. Adjacent sections of these embryos were hybridized with either the *ftz* or *h* probe alone. **a**, Dark-field image of a sagittal section of a stage 14 embryo hybridized with both *ftz* and *h* probes. Note the dorsal anterior patch of labelling, typical of *h* domain 0. The ventral surface shows a pattern of labelling which is discontinuous, the boundaries between the regions of highest labelling being indicated by the arrows. These subdivide the region into eight. Dorsally, the pattern of labelling is more uniform. Note that the posterior-most *ftz* stripe is always broader than the remainder (see, for example, Fig. 3c). **b**, **c**, **d**, Bright-field images of the ventral regions of successive sections of the embryo shown in **a**, hybridized with *ftz*, *ftz* and *h*, and *h* probes respectively. The labelled region in each case is indicated by black lines. A given cell is taken to be labelled if two or more silver grains are present in its outer cytoplasm. In **c**, arrows indicate cells adjudged to be unlabelled. The size of the labelled domains in **c** is consistent with their being due to the superposition of the individual *h* and *ftz* domains, the *ftz* domain being two nuclei anterior to the *h* domain. Note that the width of the unlabelled regions in the embryo hybridized with *ftz* never exceeds that of the regions labelled with the *h* probe.

relationship between the segmental primordia and the major axis of the blastoderm made in the fate-mapping study of Hartenstein *et al.*²⁶, but confirms the earlier observations of Lohs-Schardin *et al.*¹⁹. This asymmetry reflects a requirement for determinants of both major embryonic axes in the establishment of the *h* pattern. Our results show that in the absence of dorso-ventral polarity the *h* transcription pattern becomes approximately symmetrical. This could imply the existence of a cartesian

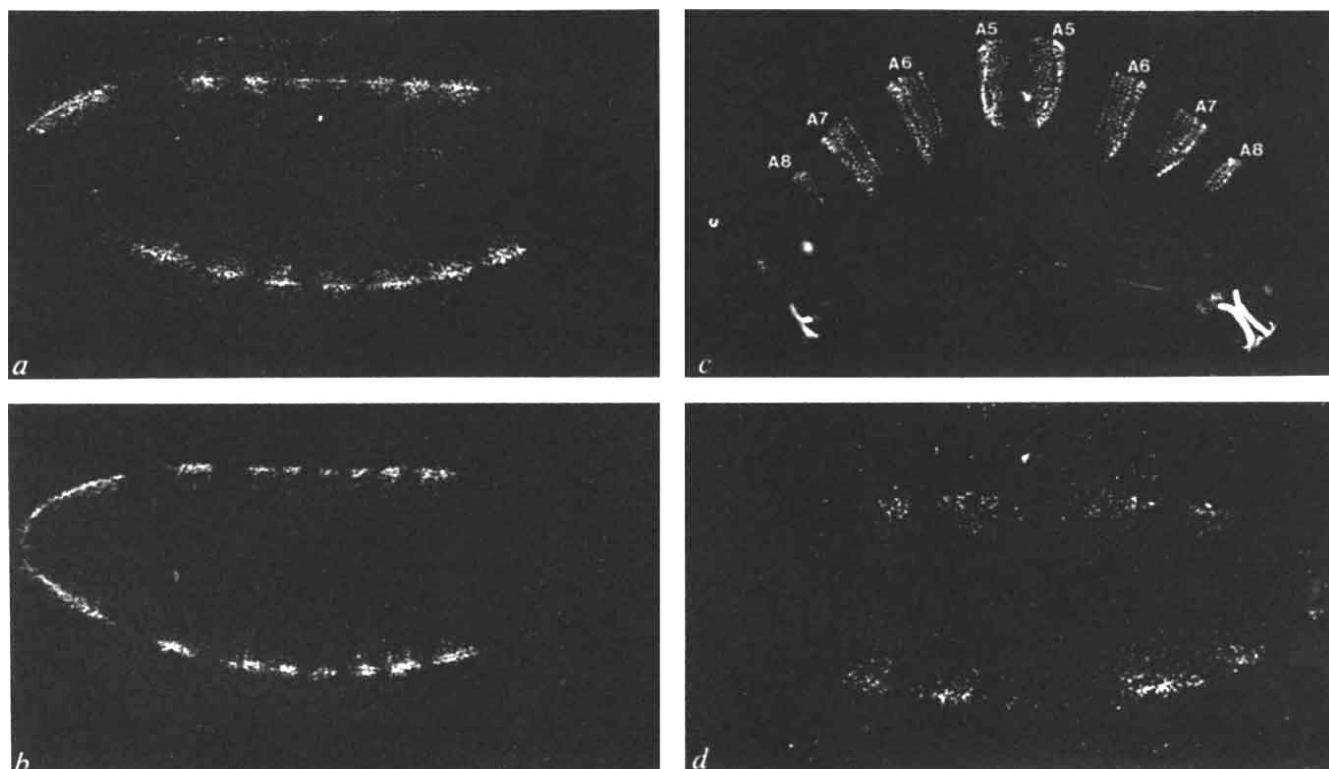


Fig. 6 *a*, Dark-field image of section shown in Fig. 4*a*, illustrating the wild-type *h* pattern. *b*, Sagittal section of a mid-stage 14 blastoderm derived from a *dl¹/dl¹* female hybridized with *h* probe. Note the accumulation of *h* transcript both dorsally and ventrally at the anterior pole. *c*, Ventral cuticular pattern of a pharate first-instar larva derived from a homozygous *bic^{M241}* female. The anterior region is replaced by segments of posterior character (A5–A8) in mirror image to the normal posterior. Note the reduction in the overall number of segments. This phenotype is variable but fully penetrant at 18 °C. *d*, Parasagittal section of a sibling of the animal shown in *c* fixed at mid-stage 14 and hybridized with *h* probe. The labelling pattern reveals a reduction in the number of regions of *h* transcript accumulation. This pattern can be interpreted as representing two regions, 6 and 7, in mirror-image symmetry, consistent with the reduction in segment number typical of these embryos (compare with Figs 1*a* and 4*a*). Note also the absence of the anterior dorsal region (0) of transcript accumulation (compared with *a* and *b*).

coordinate system, *h* responding to different antero-posterior values according to position along the dorso-ventral axis. Alternatively, the coordinate system may itself be distorted, reflecting directly the shape of the blastoderm fate map.

We thank Michael Akam and Alfonso Martinez-Arias for advice and stimulating discussions; Jonathan Slack, Jim Smith

and James Vernon for instruction in histology; Gerd Jurgens for making available unpublished data; Janni Nusslein-Volhard and Tom Kaufmann for fly stocks; Matt Scott for the *ftz* clone; Sheena Pinchin, Mimi Broit and Tracy Chaplin for technical assistance; and Rita Harris and Jyotsna Dalal for typing the manuscript.

Received 12 August; accepted 18 October 1985.

- Hogan, B., Holland, P. & Schofield, P. *Trends Genet.* **1**, 67–74 (1984).
- Sander, K. *Adv. Insect Physiol.* **12**, 125–238 (1976).
- Sander, K. in *Pattern Formation* (eds Malacinski, G. M. & Bryant, S. V.) 245–268 (Macmillan, New York, 1984).
- Lawrence, P. A. *Cell* **26**, 3–10 (1981).
- Nusslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
- Nusslein-Volhard, C., Wieschaus, E. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 267–282 (1984).
- Jurgens, G., Wieschaus, E., Nusslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 283–295 (1984).
- Wieschaus, E., Nusslein-Volhard, C. & Jurgens, G. *Wilhelm Roux Arch. dev. Biol.* **193**, 296–307 (1984).
- Wieschaus, E., Nusslein-Volhard, C. & Kluding, G. *Dev. Biol.* **104**, 172–186 (1984).
- Wieschaus, E. & Gehring, W. *Dev. Biol.* **50**, 249–263 (1976).
- Szabad, J., Schupbach, T. & Wieschaus, E. *Dev. Biol.* **73**, 265–271 (1979).
- Simcox, A. A. & Sang, J. H. *Dev. Biol.* **97**, 212–221 (1983).
- Nusslein-Volhard, C. in *Determinants of Spatial Organisation* (eds Subtelney, S. & Konigsberg, I. R.) 185–211 (Academic, New York, 1979).
- Schubiger, G., Moseley, R. C. & Wood, W. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2050–2053 (1977).
- Hafen, E., Kuroiwa, A. & Gehring, W. G. *Cell* **37**, 833–841 (1984).
- Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. *Dev. Biol.* **102**, 147–172 (1984).
- Nusslein-Volhard, C., Wieschaus, E. & Jurgens, G. *Verh. dt. zool. Ges.*, 91–104 (1982).
- Ingham, P. W., Pinchin, S. M., Howard, K. R. & Ish-Horowitz, D. *Genetics* (in the press).
- Lohs-Schardin, M., Cremer, C. & Nusslein-Volhard, C. *Dev. Biol.* **73**, 239–255 (1979).
- Jurgens, G., Lehmann, R., Schardin, M. & Nusslein-Volhard, C. *Wilhelm Roux Arch. dev. Biol.* (in the press).
- Holmgren, R. *EMBO J.* **3**, 569–573 (1984).
- Martinez-Arias, A. & Lawrence, P. A. *Nature* **313**, 639–642 (1985).
- Ish-Horowitz, D., Howard, K. R., Pinchin, S. M. & Ingham, P. W. *Cold Spring Harb. Symp. quant. Biol.* **50** (in the press).
- Foc, V. E. & Alberts, B. M. *J. Cell Sci.* **61**, 31–70 (1983).
- Akam, M. E. & Martinez-Arias, A. *EMBO J.* **4**, 1689–1700 (1985).
- Hartenstein, V., Technau, G. M. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **194**, 213–216 (1985).
- Underwood, E. M., Turner, F. R. & Mahowald, A. P. *Dev. Biol.* **74**, 286–301 (1980).
- Nusslein-Volhard, C. *Wilhelm Roux Arch. dev. Biol.* **183**, 249–268 (1977).
- Mohler, J. & Wieschaus, E. F. *Cold Spring Harb. Symp. quant. Biol.* **50** (in the press).
- Gergen, J. P., Coulter, D. & Wieschaus, E. F. *Symp. Soc. dev. Biol.* (in the press).
- Kornberg, T., Siden, I., O'Farrell, P. & Simon, M. *Cell* **40**, 45–53 (1985).
- Fjose, A., McGinnis, W. J. & Gehring, W. G. *Nature* **313**, 284–289 (1985).
- Morata, G. & Lawrence, P. A. *Nature* **255**, 614–617 (1975).
- Garcia-Bellido, A., Lawrence, P. A. & Morata, G. *Scient. Am.* **241**, 102–110 (1979).
- Kornberg, T. *Dev. Biol.* **86**, 363–372 (1981).
- Matsuda, R. *Mem. Am. Inst. Ent.* **4**, 1–334 (1965).
- Struhl, G. *Dev. Biol.* **84**, 386–396 (1981).
- Jurgens, G. *Nature* **316**, 153–155 (1985).
- Ingham, P. W., Martinez-Arias, A., Lawrence, P. A. & Howard, K. R. *Nature* **317**, 634–636 (1985).
- Struhl, G. *Nature* **308**, 454–457 (1984).
- Garcia-Bellido, A., Ripoll, P. & Morata, G. *Nature new Biol.* **245**, 251–253 (1973).
- Turner, F. R. & Mahowald, A. P. *Dev. Biol.* **57**, 401–416 (1977).
- Zalokar, M. & Erk, I. B. *J. Microbiol. Cell* **25**, 97–106 (1976).
- Zalokar, M. & Erk, I. *Stain Technol.* **52**, 89–95 (1977).
- Mitchison, T. J. & Sedat, J. *Dev. Biol.* **99**, 261–264 (1983).
- Akam, M. E. *EMBO J.* **2**, 2075–2084 (1983).
- Hafen, E., Levine, M., Garber, R. & Gehring, W. J. *EMBO J.* **2**, 617–623 (1983).
- Hayashi, S., Gillam, I. C., Delaney, A. D. & Tener, J. M. *J. Histochem. Cytochem.* **26**, 677–679 (1978).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
- Cox, K. H., DeLeon, D. V., Angerer, L. M. & Angerer, R. C. *Dev. Biol.* **101**, 485–502 (1984).
- Laughan, A. & Scott, M. P. *Nature* **310**, 25–31 (1984).
- Technau, G. M. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **194**, 196–212 (1985).

Are some BL Lac objects artefacts of gravitational lensing?

Jeremiah P. Ostriker & Mario Vietri

Princeton University Observatory, Peyton Hall, Princeton, New Jersey 08544, USA

BL Lac objects are extragalactic, highly variable, polarized sources with significant emission from radio to X-ray wavelengths. However, despite extensive research, their nature and origin remain unclear. We propose here that a significant fraction of the BL Lac objects are optically violently variable quasars (OVVs) whose continuum emission has been greatly amplified, relative to the line emission, by point-like gravitational lenses (either stars or black holes) in intervening galaxies. Several anomalous physical and statistical properties of BL Lacs can be understood on the basis of this model, which is immediately testable on the basis of absorption line studies and by direct imaging.

The class properties of BL Lacs are fairly well-defined, although this may mask an underlying heterogeneity which will retard our comprehension as it did previously with Seyferts, nebulae, or even stars. They are typically point-like (<2 arc s) with the occasional case of extended optical or radio emission perhaps critical to understanding them. In all characteristics so far described, they resemble the OVVs, a class of quasars known for high polarization and variability. They also resemble the OVVs in their optical continuum spectra, where both are steeper than typical quasars, and in their ratio of radio-to-optical where both show higher values. However, they are distinct from the OVVs and from other QSOs in the absence or near-absence of the normal QSO emission lines, even though, during low periods of continuum, emission lines are occasionally seen¹. Statistically, BL Lac objects are uncommon among the brighter extragalactic point sources, constituting perhaps 10% of all objects brighter than $m_B = 16$ (~ 35 objects over the whole sky), and even less common at fainter magnitudes². This is also true of their X-ray properties, where the number-magnitude relation flattens out much earlier than it does for normal QSOs³.

The conventional interpretation is that BL Lac objects are one type of active galactic nuclei. This is supported by the fact that several have been found in relatively small-redshift elliptical galaxies⁴. The variability, high polarization and absence of extended radio structure are attributed to relativistic beaming⁵. While few detailed models have been constructed⁶, the picture is moderately persuasive and is likely to be correct for some subset of the BL Lac objects. There are, however, seriously anomalous features of the observations which are inconsistent with or unexplained by the standard interpretation.

First, there are physical problems. The emitted ionizing continuum would be expected to be absorbed by ambient gas in the elliptical galaxy, with some radiation ultimately degraded to the H α line. Using standard results⁷, we find that if the gas in the elliptical satisfies the inequality $n_e M_{\text{gas}} > 10^{8.5} M_\odot \times \text{atoms cm}^{-3}$, then H α would be detectable in a $m_B = 16$ BL Lac at $z = 0.05$. This is comparable with the amount of gas found in ellipticals, and so, if the gas has not been ejected due to violent nuclear activity, might be expected to be seen in some cases.

Then come the problems with the popular beaming model: it is unlikely that it can, alone, account for the properties of compact radio sources. If the sources whose beams are pointed away from our line-of-sight are identified with radio-quiet quasars, which are otherwise identical to compact sources, why do the compact sources have isotropically emitting haloes, while the extended sources do not? If, on the other hand, extended steep spectrum sources are to be identified with the sources beamed away from us, then the statistics fail, in the sense that there are too few extended QSOs to be consistent with a random-pointing beam. Also, there is at least one extended source

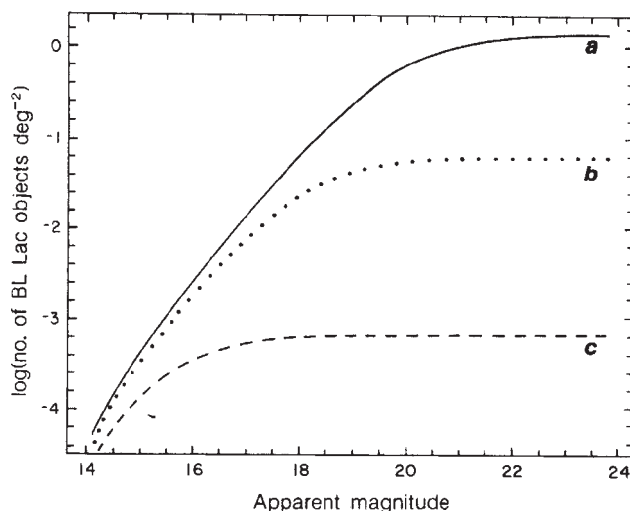


Fig. 1 Number-magnitude relation for OVVs that have been amplified by mini-lensing by at least a factor of 2(a), 10(b) or 100(c), per unit square degree. The OVVs are supposed to have the same luminosity function as the QSOs¹⁶, and to constitute only a fraction $f = 10\%$ of all QSOs.

(3C179) that shows superluminal expansion, and one compact source (4C39.75) that does not.

There are also statistical incongruities. If BL Lacs are nearby objects, a euclidian number-flux relation in the X-ray $N \propto f_x^{-1.5}$ would be expected which is significantly steeper than observed³ for $f_x < 10^{-12}$ erg s⁻¹ cm⁻². This result is not biased by the usual incompleteness of the optical data. Then there is the distribution in distance. Only a small number of BL Lac objects have been examined for absorption line redshifts, with identified systems ranging from $z_{\text{abs}} = 0.424$ (PKS0753+178)⁸ to $z_{\text{abs}} = 1.719$ (0215+15)⁹. This is not unlike the QSO distribution, but it makes anomalous the relatively large number of objects apparently seen in nearby ($z < 0.2$) bright ellipticals. If the typical QSO has $z = 1-1.5$, then the nearest of a sample of 60 should have a $z = 0.16$ (for $\Omega = 1$). This argument is valid for QSOs where the nearest of the hundred brightest objects is 3C273 with $z = 0.158$. But it is highly improbable ($< 10^{-2}$ to find a single object) to have a fraction $\sim 50\%$ of the objects with known redshifts such that $z < 0.2$ as is apparently the case for the BL Lacs. One could avoid this dilemma by imagining the number-luminosity relation to be very steep with many more intrinsically faint than bright objects in comparison with normal quasars, but then the relatively flat number-magnitude relation² becomes even more peculiar. But, if we treat the nearby BL Lacs apparently in galaxies as statistical accidents to avoid these difficulties, we are faced with a wild improbability: one quasar¹⁰ out of 2,000 has been found with its image superimposed on the central regions of a galaxy having $z < 0.2$; so 11 of the 50-odd BL Lacs is extremely unlikely on statistical grounds.

On the basis of this argument, it is attractive to look for an explanation for the phenomenon in which the BL Lacs are at cosmological distances, but the apparently foreground galaxies, whether seen in absorption or in emission, are not there accidentally. We propose that a significant fraction have been gravitationally lensed by point-like objects in the foreground galaxies. Typical optical depths for mini-lensing¹¹ are of the order of 10% for $\Omega = 0.2$, if the haloes of galaxies are made of stellar-like objects. Thus we propose that the OVVs are intrinsically small (or 'beamed') objects, on the basis of the usual variability arguments, and that of the order of 10% of them have been magnified by mini-lenses in intervening galaxies to present us with a BL Lac object. This is made possible by the differing sizes of the continuum-emitting regions ($r \sim 10^{15}$ cm) and of the line-emitting regions ($r \sim 0.1-1$ pc). The line-emitting regions are too large to be lensed by stars, or even black holes up to

the mass of 10^1 – $10^3 M_\odot$ (ref. 12). This model is similar to that proposed by Barnothy and Barnothy¹³, but we are applying it here to only a small subset of QSOs. Normal QSOs are too large, in general, to be affected by mini-lensing.

In recent studies of the effects of lensing on the statistics of QSOs^{14,15}, we found that a class of objects with the postulated properties should be expected to exist on statistical grounds. We have computed the number-magnitude relation for all OVVs that have been brightened by mini-lensing by at least a factor of 2, 10 or 100; the adopted luminosity function is that given by Schmidt and Green¹⁶, truncated at $z = 3.5$, and scaled down by a factor of 10, to account roughly for the incidence of OVVs among regular quasars. The results are shown in Fig. 1. Given the prominence of the emission lines of QSOs, it is likely that only objects whose continuum has been brightened by at least a factor of 10 would be classified as BL Lacs. Figure 1 shows the expected apparent magnitude distribution for such objects, assuming a density of point-like lenses equivalent to $\Omega_l = 0.2$ (refs 14, 15). The striking feature is the high apparent luminosity at which the counts flatten: $m_B = 16$ –18. This agrees qualitatively with the X-ray observations, and explains the negative results of optical searches for faint ($m_B \sim 20$) polarized objects²: both surveys investigated only a small area of sky, where but a handful of objects would be found on the basis of the statistics of regular quasars. On the contrary, we predict that searches would be most successful for large areas of sky and relatively bright limiting magnitudes. Our computation predicts ~ 100 BL Lacs over the whole sky, in agreement with the 50-odd objects discovered in (mostly) the Northern Hemisphere.

The redshift distribution of BL Lacs would be, of course, an insurmountable problem for our model if we accepted that the redshifts attributed to them are their own. But in our model it is natural to suppose that the stellar spectra seen in either emission or absorption are due to lensing galaxies, which are not physically associated with the BL Lacs; the latter are instead at cosmological distances. The redshift distribution expected for BL Lacs, for limiting magnitude of 16, is shown in Fig. 2; it mimics that of QSOs quite closely¹⁶.

There remains an apparent discrepancy for this picture in the comparison between predicted and observed redshift distributions for the lensing galaxies, which lie relatively close to us. Models¹¹ (especially Fig. 2 of ref. 11) show that they should primarily have $z \sim 0.2$ – 0.7 , if they are uniformly distributed in space ($n_l = n_0(1+z)^3$), while we know of at least eight objects¹⁷ which show absorption lines due to stellar spectra, all with $z < 0.2$. But this problem arises also when one considers normal QSOs and lensing by a galaxy, as opposed to mini-lenses; of the six known lensed quasars, one¹⁰ is lensed by a low-redshift galaxy. A tentative explanation for this may be offered (B. Paczynski, personal communication). In accounting for the large splitting and amplification of the known multiple quasars, it was pointed out¹¹ that, if the galaxies are in a cluster with surface mass density Σ approaching the critical surface mass density Σ_{crit} , then the flux is increased by $(1 - \Sigma/\Sigma_{crit})^{-2}$. Similarly, although the analogy is imperfect because the mass distribution in the lensing galaxy will be very grainy, we may expect mini-lensing to be most effective if the stars are located in a galaxy having a mean surface mass density close to the critical density. We find (equation 2.35 of ref. 11) that $\Sigma_{crit} = H_0 c / (4\pi z_l G) = 0.12 h/z_l \text{ g cm}^{-2}$, for $z_l \ll z_{QSO}$, regardless of cosmology. But the mean surface density within the de Vaucouleurs radius is relatively independent of luminosity for elliptical galaxies and is typically of the order of $(1\text{--}2)h \text{ g cm}^{-2}$ (ref. 18). Thus, normal ellipticals are near the critical surface mass density for redshifts of the order of $z \sim 0.05$ – 0.1 , which will make mini-lensing in this redshift range more prominent due to the 'amplification bias' already noted¹¹. As the lensing properties approach those for continuum mass distributions as $\Sigma \rightarrow \Sigma_{crit}$, we would expect some BL Lacs in this redshift range to show multiple images, as is the case for the object observed by Huchra *et al.*¹⁰.

Monte Carlo simulations of flux variations due to mini-lensing¹⁹ for 0957+561B show that an average amplification

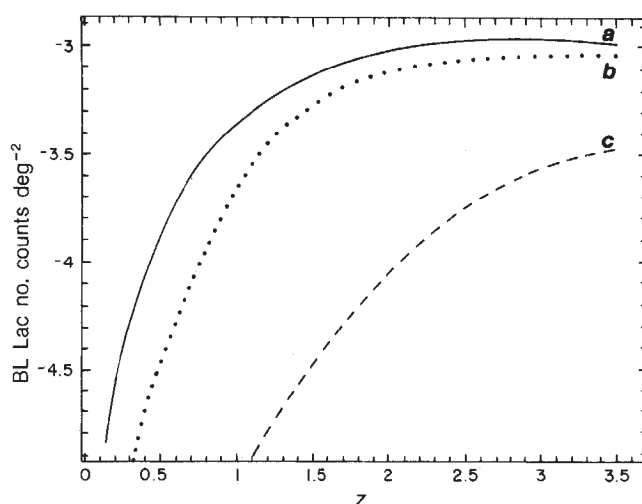


Fig. 2 Redshift distribution for all OVVs with $m_B \leq 16$, which mini-lensing has amplified by at least a factor of 2(a), 10(b) or 100(c), per unit redshift.

$A \sim 8$ can be obtained for a sufficiently small quasar; the time-scale for flux variation is ~ 5 yr for solar mass stars, a value not inconsistent with that of the observed large-amplitude optical fluctuations. Note also that the relatively small amplitude of the fluctuations obtained by Young¹⁹ (a factor of ~ 2) may be due to the large beam size assumed for 0957+561, which is not an OVV. Also, large swings in the polarization vector during outbursts may be due to mini-lensing²⁰: when the optical depth to mini-lensing is $\tau_{ml} \sim 1$, any subregion of the quasar image can be amplified; which region actually is, depends on the very non-local cooperation of all mini-lenses. If the polarization changes intrinsically in the source, the star motions cause differently polarized regions to be amplified at different times, and the polarization vector to swing during outbursts.

We emphasize that mini-lensing is not alternative to beaming, but depends on it, in that mini-lensing can cause some beamed sources (OVVs) to become BL Lacs. In fact, mini-lensing can relieve the statistical problems of the beaming hypothesis. Tests comparing local densities of BL Lacs and of QSOs have concluded¹⁷ that it is difficult to find a class of radio sources which is numerous enough to be consistent with the statistics of a randomly pointing small beam. If our hypothesis is correct, though, the BL Lacs are at cosmological distances, and this implies that the local density of BL Lacs has been overestimated¹⁷.

The last objection to the mini-lenses hypothesis that we consider here is the following: because we assume that mini-lensing can amplify the optical continuum, but not the line-emitting region, the radio continuum ($r \sim 1$ pc) should not be amplified, and OVVs and BL Lacs should have significantly different (by a factor of 10) radio-to-optical ratios. This objection neglects K-corrections, which are necessary in our hypothesis because the BL Lacs are at cosmological distances. The radio continuum is rather flat ($\alpha_r \sim 0$), while the optical continuum is steep ($\alpha_{opt} \sim 2$), so neglect of K-corrections (of order $(1+z)^2 \sim 10$ for cosmological redshifts) can easily mask different radio/optical flux ratios.

Our simple model has several consequences which make it easy to test.

- (1) One should expect to see absorption redshift systems in a significant fraction of BL Lacs, as they lie at cosmological distances, especially in the objects that show absorption due to stellar spectra. In these objects, one expects to have $z_{abs} > z_{gal}$. As it has been argued that most quasars have $z \leq 2$, we also expect this to be the case for BL Lacs, as observed²¹.
- (2) We predict a continuity of properties between BL Lacs and OVVs, mostly in the relative importance of the line and of the

continuum emission. This continuity has been noticed already for several other properties (polarization, fluctuations of flux and polarization and their timescales)²², with the BL Lacs always having the most extreme properties.

(3) The position of the optical continuum source should not be centred on the image of the associated galaxy, where observed. Two examples of this sort^{1,23}, where the distances are 6.5 and 12 arc s, respectively, are known already, whereas if the BL Lac object is in the galaxy, it should always be centrally located. One might also expect a few objects to have multiple images, or the radio source not to coincide with the optical galaxy²⁴.

(4) But if, contrary to our hypothesis, the strong hard spectrum source sits in the core of a nearby elliptical galaxy, one would often expect to see Balmer line emission from the galaxy, while in our model this should not happen, as there is no physical connection between the hard spectrum source and the galaxy.

We thank M. J. Rees, B. Paczynski and E. L. Turner for helpful discussions. This work was supported in part by NASA grant NAGW-765.

Received 3 July; accepted 13 September 1985.

1. Arp, H., Sargent, W. L. W., Willis, A. G. & Oosterbaan, C. E. *Astrophys. J.* **230**, 68–78 (1979).
2. Borra, E. F. & Coriveau, G. *Astrophys. J.* **276**, 449–453 (1984).
3. Maccacaro, T., Gioia, I. M., Maccagni, D. & Stoeck, J. T. *Astrophys. J. Lett.* **284**, L23–28 (1984).
4. Miller, J. S., French, H. B. & Hawley, S. A. in *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 176–187 (Pittsburgh University Press, 1978).
5. Blandford, R. D. & Konigl, A. *Astrophys. J.* **232**, 34–48 (1979).
6. Konigl, A. *Astrophys. J.* **243**, 700–709 (1981).
7. Osterbrock, D. in *Active Galactic Nuclei* (eds Hazard, C. & Mitton, S.) 25–50 (Cambridge University Press, 1977).
8. Carswell, R. F., Coleman, G., Strittmatter, P. A. & Williams, R. E. *Astr. J.* **53**, 275–281 (1976).
9. Blades, J. C., Hunstead, R. W., Murdoch, H. S. & Pettini, M. *Astrophys. J.* **288**, 580–594 (1985).
10. Huchra, J. et al. *Astr. J.* **90**, 691–696 (1985).
11. Turner, E. L., Ostriker, J. P. & Gott, J. R. *Astrophys. J.* **284**, 1–22 (1984).
12. Gott, J. R. *Astrophys. J.* **243**, 140–146 (1981).
13. Barnothy, J. M. & Barnothy, M. F. *Science* **162**, 248–250 (1968).
14. Vietri, M. *Astrophys. J.* **293**, 343–355 (1985).
15. Ostriker, J. P. & Vietri, M. *Astrophys. J.* (in the press).
16. Schmidt, M. & Green, R. F. *Astrophys. J.* **269**, 352–374 (1983).
17. Schwarz, D. A. & Ku, W. H.-M. *Astrophys. J.* **266**, 459–465 (1983).
18. Kormendy, J. *Astrophys. J.* **218**, 333–354 (1977).
19. Young, P. J. *Astrophys. J.* **244**, 756–767 (1981).
20. Paczynski, B. Princeton University Observatory preprint 135 (1985).
21. Peterson, B. M. *Astrophys. Lett.* **20**, 119–122 (1980).
22. Angel, J. R. P. & Stockman, H. S. *A. Rev. Astr. Astrophys.* **18**, 321–362 (1980).
23. Craine, E. R. & Warner, J. W. *Astrophys. J.* **206**, 359–363 (1976).
24. Danziger, J. S., Fosbury, R. A. E., Goss, W. M. & Ekers, R. *Mon. Not. R. astr. Soc.* **188**, 415–419 (1979).

GB841215, the fastest γ -ray burst?

J. G. Laros*, E. E. Fenimore*, M. M. Fikani*,
R. W. Klebesadel*, M. van der Klis† & M. Gottwald‡

* Los Alamos National Laboratory, Los Alamos,
New Mexico 87545, USA

† Space Science Department of ESA, ESTEC, Postbus 299,
2200 AG Noordwijk, The Netherlands

‡ EXOSAT Observatory, Space Science Department of ESA, ESOC,
Darmstadt, FRG

In the 12 yr since the discovery of γ -ray bursts by Klebesadel *et al.*¹, several hundred of these enigmatic events have been observed and catalogued (see, for example, refs 2–5). Their time histories have exhibited a tremendous diversity: they can have durations of milliseconds or minutes; they may contain one or a dozen individual peaks; and they can be highly impulsive or slowly varying. Possibly the only truly unifying observational property of γ -ray bursts is that, with very few exceptions^{6,7}, the bulk of the energy output seems to be in the form of γ -rays. Here we report the detection on 15 December 1984 at 08.25 UT of an extraordinary outburst, qualitatively different in appearance from all previously observed γ -ray bursts. It is described most conveniently as a 'classical' multi-peaked, hard-spectrum burst that has been compressed in time by a factor of 10–100 while simultaneously having its intensity increased by a like factor (thus conserving fluence). Its peak intensity was much higher than any other known γ -ray burst except

for GB790305b, which had an unusually soft spectrum and other unique features that set it apart from 'classical' γ -ray bursts⁸. However, a key point is that if the intensity of GB841215 had been 'normal', its narrow individual spikes would not have been statistically significant with current instrumentation, and the event would have had the appearance of a rather ordinary, short γ -ray burst.

GB841215 was observed by the Los Alamos γ -Burst Detector experiment on the Pioneer Venus Orbiter (PVO)⁹, by the UCB/Los Alamos Solar X-ray/ γ -Ray Burst experiment on the International Cometary Explorer (ICE, previously ISEE 3)¹⁰, and by the medium-energy detectors¹¹ on the European Space Agency's X-ray observatory EXOSAT. Figure 1 shows the time history of GB841215 as obtained by the ICE and PVO experiments. Due to excessively high counting rates induced by this event, EXOSAT suffered a 'crash' of its onboard computer and lost all useful data except for an approximate event onset time. The time bins in Fig. 1 are of unequal length because both the ICE and PVO experiments store data in memories that automatically convert to a 'time-to-spill' mode whenever the rate exceeds $1,365 \text{ s}^{-1}$, that is, the higher rates are derived from the time interval required to accumulate a specific number of counts (16 in Fig. 1a and 32 in Fig. 1b). The statistical properties of such a plot are quite different from those of the familiar Poisson distribution and tend to produce a spikier appearance because the fractional standard deviations do not decrease with increasing rate. The rate changes seen in Fig. 1a, b are significant if they exceed a factor of ~ 2 or ~ 1.4 , respectively. Thus, during the ~ 0.3 -s impulsive phase of the event, at least seven distinct peaks are evident in both parts of Fig. 1, with the sharpest of these having a full width of $\leq 0.005 \text{ s}$ in Fig. 1a. No other γ -ray burst has exhibited such rapid multi-peaked structure (hence the word 'fastest' in the title, even though GB790305b may have had a faster rise time). However, it would be unwarranted to conclude that such narrow multiple spikes are rare or unique, because the ability to detect rapid variations is a strong function of intensity—especially in systems with rate-dependent time resolution. If GB841215 had been a factor of 10 or more weaker, it would have had peak rates comparable with other intense bursts, but, its remarkable signature would not have been discernible. Instead, a rather uninspiring 0.3-s event with probable ~ 0.1 -s structure would have been observed. (In fact, other short events observed by PVO have had marginally significant—but never convincing—indications of very fast structure.)

Because of the different energy thresholds of the two experiments (260-keV minimum for the ICE memory data versus 100 keV for PVO), the respective time histories are qualitatively different. In Fig. 1a (higher threshold), the peaks are sharper and the 'valleys' are broader and deeper, relative to b. Evident only in Fig. 1b is the beginning of a weak tail with a softer spectrum that is present for $\sim 10 \text{ s}$. The tail does not show any clear modulations, but we have not yet thoroughly investigated this matter. The spectrum of GB841215 is generally very hard, with considerable emission above 1.5 MeV. The total >30 -keV fluence of $4 \times 10^{-4} \text{ erg cm}^{-2}$ is large, but not exceptional. The peak intensity of $\sim 2.5 \times 10^{-3} \text{ erg cm}^{-2} \text{ s}^{-1}$ is exceptional, and has been exceeded only by GB790305b.

The location information for GB841215 is derived from only a three-satellite arrival-time analysis. Accordingly, there are two non-redundantly determined error boxes located approximately symmetrically with respect to the ecliptic plane. (We are still searching for additional data that could lead to a better localization for this event, but the probability of success is not high.) The $\geq 90\%$ confidence error boxes are centred at $(\alpha, \delta)_{1950} = (16 \text{ h } 44 \text{ min } 48.2 \text{ s}, -6^{\circ}42'59'')$ and $(16 \text{ h } 17 \text{ min } 09.2 \text{ s}, -42^{\circ}24'10'')$, or $(L_2, B_2) = (11.2, 23.6)$ and $(338.8, 5.3)$. They are best described as 2° -long segments of an annulus whose axis is located at $(\alpha, \delta)_{1950} = (21 \text{ h } 13 \text{ min } 19.4 \text{ s}, -18^{\circ}10'20'')$ and whose radius is $66^{\circ}13'19'' \pm 4''$. The ends of the two segments are located at $(16 \text{ h } 43 \text{ min } 37.9 \text{ s}, -7^{\circ}40'57'')$, $(16 \text{ h } 45 \text{ min } 56.3 \text{ s}, -5^{\circ}48'03'')$ and $(16 \text{ h } 17 \text{ min } 29.6 \text{ s}, -41^{\circ}23'44'')$, $(16 \text{ h } 16 \text{ min } 51.7 \text{ s}, -43^{\circ}21'33'')$. The total error box area is 33 arc min^2 . Searches

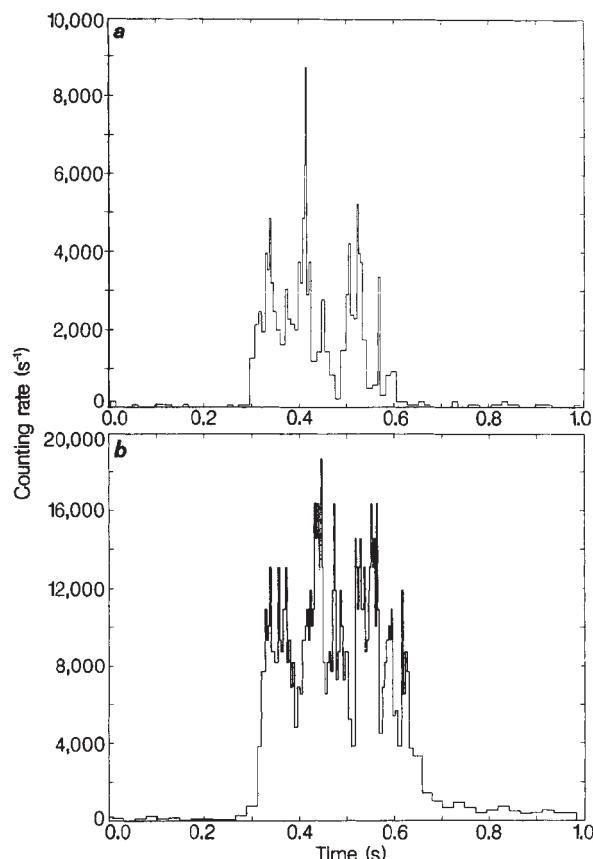


Fig. 1 *a*, ICE time history of GB841215. The energy range is 0.26–2.5 MeV, and the time resolution varies from 0.002 s to 0.012 s. *b*, PVO time history of GB841215. The energy range is 0.1–2.0 MeV, and the time resolution varies from 0.0017 s to 0.012 s.

through various catalogues of γ -ray bursts, pulsars, supernova remnants, flare stars, and high-energy emitters did not result in any source candidates.

We have observed a γ -ray burst with an apparently unique time history. The uniqueness stems from the event's intensity and the rapidity of its temporal structure. However, it is not clear whether the time history is really unique, or whether other events have had similar structure that went undetected due to lack of sufficient event intensity and/or experiment time resolution. If the latter is true, then we must ask if the true 'face' of a γ -ray burst has ever been seen. Are any of our observed γ -ray burst peaks really single peaks? Have any of our measurements (including this one) revealed the underlying fluctuation time scale of a γ -ray burst? Are measured γ -ray burst spectra distorted by pulse pile-up effects? (This could be the case if the photons are emitted preferentially in packets only moderately shorter than the peaks observed in GB841215.) A more detailed discussion of these and other points, as well as a detailed spectral and temporal analysis of this event, will be presented elsewhere.

We thank N. White for valuable help in expediting communications between the widely separated institutions that participated in this research. This work was supported by the US Department of Energy and by NASA under contracts A-98331A and NAS 5-25980.

Received 17 June; accepted 19 September 1985.

1. Klebesadel, R. W., Strong, I. B. & Olson, R. A. *Astrophys. J. Lett.* **182**, L85–L88 (1973).
2. Mazets, E. P. *et al. Astrophys. Space Sci.* **80**, 3–143 (1981).
3. Klebesadel, R. W. *et al. Astrophys. J. Lett.* **259**, L51–L56 (1982).
4. Baity, W. A., Hueter, G. J. & Lingenfelter, R. E. *AIP Conf. Proc.* No. 115, 434–484 (1984).
5. Atteia, J.-L. *et al. Astrophys. J. Suppl.* (in the press).
6. Mazets, E. P., Golenetskii, S. V., Guryan, Yu. A. & Ilyinski, V. N. *Astrophys. Space Sci.* **84**, 173–189 (1982).
7. Laros, J. G., Fenimore, E. E., Klebesadel, R. W. & Kane, S. R. *Bull. Am. astr. Soc.* **17**, 520–521 (1985).
8. Cline, T. L. *et al. Astrophys. J. Lett.* **237**, L1–L6 (1980).
9. Klebesadel, R. W. *et al. IEEE Trans. Geosci. Rem. Sens.* **GE-18**, 76–80 (1980).
10. Anderson, K. A. *et al. IEEE Trans. Geosci. Elect.* **GE-16**, 157–159 (1978).
11. Turner, M. J. L., Smith, A. & Zimmerman, H. U. *Space Sci. Rev.* **30**, 513–524 (1981).

Change of solar oscillation eigenfrequencies with the solar cycle

Martin F. Woodard* & Robert W. Noyes

Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA

*Present address: Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

Solar acoustic eigenfrequencies depend on the internal structure of the Sun, which may change during the 11-yr cycle of magnetic activity as a result of various effects associated with the solar dynamo. Observations of low-degree acoustic frequencies were made, using the ACRIM instrument on the Solar Maximum Mission (SMM) satellite, in 1980 (near solar maximum) and 1984 (near solar minimum). The analysis of these data, presented here, indicates that the frequencies of $l=0$ and $l=1$ acoustic modes in the 5-min band have decreased from 1980 to 1984, by $\sim 0.42 \mu\text{Hz}$ or 1.3 parts in 10^4 . This finding may have important implications for our understanding of the mechanism of the solar activity cycle.

The ACRIM solar total irradiance data have been described in detail elsewhere¹. For the analysis of 5-min oscillations we use a series of flux estimates obtained at the rate of $1/131.072$ s, as defined by the periodic chopping of the solar input to the ACRIM sensor by mechanical shuttering. These data are interrupted by the passage of the SMM into the Earth's shadow about every 96 min and by other, apparently random, data gaps.

Data of sufficient quality for studying the 5-min acoustic (p -mode) frequencies were obtained during 1980 from 18 February to 1 December and in 1984 from 1 May (following the successful repair of SMM) to 31 December. Data obtained in 1985 have not yet been analysed.

Frequencies of low-degree p -modes were derived for both the 1980 and 1984 epochs, by means of discrete Fourier transforms of the data from each epoch. The smoothed periodogram of the 1984 data, for example, is shown in Fig. 1. Figure 2 shows the frequency difference (1980 minus 1984) for the nine strongest oscillation peaks in the ACRIM power spectrum. With one exception, the 1980 frequencies of these modes are higher than their 1984 frequencies. To quantify this trend, we assume that all true frequency differences are identical, and that the error distribution is the same for all measured modes. Then we calculate the mean frequency difference, and estimate its error from the standard deviation of the mean of the points:

$$\Delta\nu = 0.42 \pm 0.14 \mu\text{Hz} \quad (1)$$

The 1980 and 1984 observing windows differ in their details. In particular, data from days 243–273 of 1984 were not available

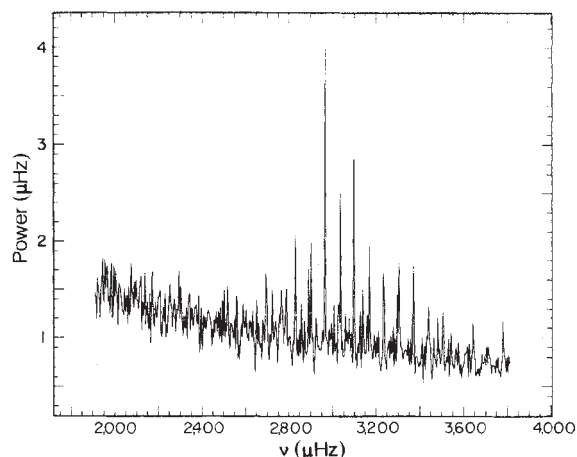


Fig. 1 Smoothed periodogram of 1984 ACRIM data in the 5-min band, showing solar oscillation modes of degree $l=0, 1$ and 2, and radial order n in the range ~ 17 –26. The unit of power is mean square fractional irradiance variation.

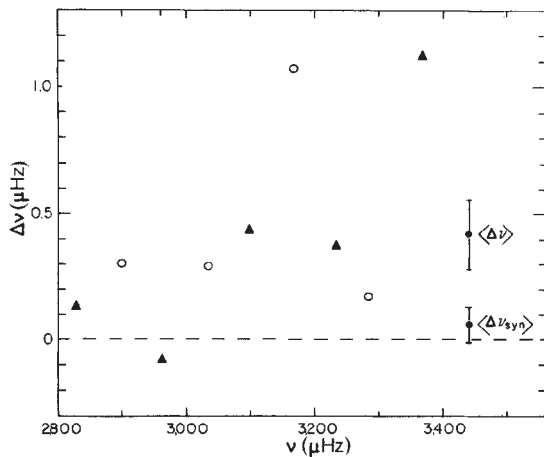


Fig. 2 Measured frequency differences (1980 minus 1984) for individual 5-min solar acoustic modes of degree $l=0$ (○) and $l=1$ (▲), plotted against the frequencies of the modes. The mean and standard deviation of the mean of these nine measurements are indicated by the point labelled $\langle \Delta \nu \rangle$ (equation (1)). The estimated frequency pulling due to the different observing windows (1980 minus 1984) is shown by the point labelled $\langle \Delta \nu_{\text{syn}} \rangle$ (equation (2)).

at the time of the analysis. To test for possible systematic frequency pulling due to the different window functions, we manufactured and analysed time strings sharing properties with the real data. Each synthetic time string is the superposition of variations due to 75 individual oscillation modes plus random variations representing noise. The variation corresponding to an individual mode is patterned after the response of a damped simple harmonic oscillator to a random driving function. The amplitude of the variations representing both those modes seen in the ACRIM data set and random noise was chosen to match the corresponding observed amplitudes in that data set. The amplitude and frequency of the remaining modes, not detected in the ACRIM data, were estimated from velocity oscillation data².

Five independent strings, or runs, of synthetic data were generated. Each run is a separate realization of a statistical process, defined by parameters such as resonant frequencies, linewidths and amplitudes, whose values are the same for all runs. Each randomly generated time series was masked by both the 1980 and the 1984 observing windows and the resulting pair of intermittent data sets from each run was analysed in the same manner as the real data. We therefore obtained five independent estimates of possible frequency shifts due to differences in the window function. From the mean and standard deviation of the mean of five measured shifts, we obtain an estimate of this systematic frequency shift

$$\Delta \nu_{\text{syn}} = 0.06 \pm 0.07 \mu\text{Hz} \quad (2)$$

consistent with no systematic shift. We may obtain a more conservative estimate of the frequency shift of the solar oscillations by subtracting the estimates (1) and (2) and combining their errors quadratically, to give

$$\Delta \nu = 0.36 \pm 0.15 \mu\text{Hz} \quad (3)$$

We have also checked the consistency of the frequencies by comparing portions of the 1984 time series with the whole of the 1980 data. Portion *a* consists of approximately the first 100 days of available 1984 data and portion *b* approximately the final 100 days of that year. The measured frequency shifts (with notation identical to that of equation (1)) are:

$$\Delta \nu_a = 0.43 \pm 0.19 \mu\text{Hz} \quad (4a)$$

$$\Delta \nu_b = 0.36 \pm 0.13 \mu\text{Hz} \quad (4b)$$

Thus the magnitude and sign of the frequency change is consistent between the two portions. We conclude that in 1980, near solar maximum, the low-degree *p*-mode eigenfrequencies were

systematically $\sim 0.4 \mu\text{Hz}$ higher than in 1984 near solar minimum. This result is significant at the $2.5\text{--}3\sigma$ level. We note that a similar change for the most prominent $l=1$ mode of velocity oscillations has been reported³; preliminary analysis indicated a decrease of $0.9 \mu\text{Hz}$ for that mode between 1981 and 1983. In addition, an increase of $1 \mu\text{Hz}$ was suggested for $l=0$ (but not $l=1$) modes from 1980 to 1981.

According to asymptotic theory⁴, the frequency of low-degree acoustic modes of high radial order is proportional, to leading order, to

$$\nu_0 = \left[2 \int_0^R \frac{dr}{c(r)} \right]^{-1}, \quad (5)$$

where $c(r)$ is the adiabatic sound speed and R is the radius of the upper reflection point of the waves and is close to the radius of the photosphere. Thus ν_0 is essentially the reciprocal of the sound travel time across the Sun. The modes used in our analysis (see Fig. 2) have an average frequency of $\sim 3,100 \mu\text{Hz}$. Thus, the frequency change $\Delta \nu$ implies that in 1980, near solar maximum, the sound travel time was shorter by ~ 1.3 parts in 10^4 than in 1984, near solar minimum. A plausible explanation⁵ for this result is that, to an order of magnitude, a similar fractional change in solar radius has occurred, so that the radius is some 10^2 km smaller at solar maximum than at solar minimum. There is other evidence for activity cycle-related changes in the solar radius, from historical meridian circle observations, transits of Mercury and eclipse timings⁶, this evidence also suggests a smaller solar radius at activity maximum, by about the same amount.

It is tempting to associate the observed change in solar *p*-mode eigenfrequencies with a restructuring of the outer convective envelope of the Sun accompanying the activity cycle, for changes in the observed properties of the convective zone have also been reported in association with activity changes. Specifically, both the small-scale granulation⁷ and the larger-scale supergranulation cells⁸ have been found to be significantly smaller at sunspot maximum than at sunspot minimum. Some theoretical studies⁹ have suggested that dynamo generation of magnetic fields deep in the convection zone could lead to a decreased efficiency of convection. This in turn could produce a decrease in the radius of the Sun and a slight increase in *p*-mode eigenfrequencies¹⁰.

Other effects have been suggested which might cause changes in acoustic mode eigenfrequencies correlated with the activity cycle, above and beyond any changes directly coupled to radius changes. Thus Brown¹¹ notes that, in general, turbulent velocities in the convection zone will decrease acoustic mode eigenfrequencies; a decrease in the amplitude of convective turbulence near solar maximum would then lead to an increased frequency at that time, as observed. Bogdan and Zweibel¹² point out that fibril magnetic fields in the upper convection zone, which should be more prevalent near solar activity maximum, would increase the eigenfrequencies, again as is observed. It appears that detailed modelling of a variety of effects may be necessary to explain the observed behaviour; it is hoped that such investigations will produce new insights into the mechanisms of the solar cycle.

We thank T. M. Brown, D. O. Gough, R. Rosner and C. A. Whitney for useful discussions. The reduced ACRIM data were kindly provided by R. C. Willson. This research was supported by NASA contract NAG-5-506.

Received 11 July; accepted 19 September 1985.

- Willson, R. C. & Hudson, H. S. *Astrophys. J. Lett.* **244**, L185–189 (1981).
- Grec, G., Fossat, E. & Pomerantz, M. *Sol. Phys.* **82**, 55–66 (1983).
- Van der Raay, H. B. in *Theoretical Problems in Stellar Stability and Oscillations*, 215–219 (University of Liege, 1984).
- Tassoul, M. *Astrophys. J. Suppl.* **43**, 469–490 (1980).
- Christensen-Dalsgaard, J. in *Space Research Prospects in Stellar Activity and Variability* (eds Mangenay, P. & Praderie, F.) 11–45 (Observatoire de Paris, Meudon, 1984).
- Gilliland, R. L. *Astrophys. J.* **248**, 1144–1155 (1981).
- Macris, C. J., Muller, R. Rösch, J. & Rouddier, T. in *Small-Scale Dynamical Processes in Quiet Stellar Atmospheres* (ed. Keil, S. L.) 265–295 (Sacramento Peak, Sunspot, 1984).
- Singh, J. & Bappu, M. K. V. *Sol. Phys.* **71**, 161–168 (1981).
- Spiegel, E. A. & Weiss, N. O. *Nature* **287**, 616–617 (1980).
- Gough, D. O. *NASA Conf. Publ.* No. 2191, 185–205 (1981).
- Brown, T. M. *Science* **226**, 687–688 (1984).
- Bogdan, T. J. & Zweibel, E. G. *Astrophys. J.* (submitted).

Ionospheric response to changes in the interplanetary magnetic field observed by EISCAT and AMPTE-UKS

H. Rishbeth*, P. R. Smith†, S. W. H. Cowley†,
D. M. Willis*, A. P. van Eyken*,
B. J. I. Bromage* & S. R. Crothers*

* Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK

† Blackett Laboratory, Imperial College, London SW7 2BZ, UK

A major question in solar-terrestrial physics concerns the nature of the interaction between the solar wind and the Earth's magnetosphere and ionosphere, and how this depends on the direction of the interplanetary magnetic field (IMF). We report here simultaneous observations of the IMF by the AMPTE-UKS (Active Magnetospheric Particle Tracer Explorers-UK Satellite), and of plasma flow in the high-latitude ionosphere by the EISCAT radar with better time resolution than has been achieved hitherto. We discuss one event in which a sharp southwards turning of the IMF is followed by the onset of rapid ionospheric plasma flow near local noon, demonstrating the close control of ionospheric flows by the IMF. We discuss the causes of the observed 13-min delay, and study the ionospheric heating that accompanies the acceleration.

The AMPTE-UKS¹ has a perigee of 550 km, apogee of 18.7 Earth radii and period of 43.8 h. In October 1984, apogee was near the noon meridian and the satellite was ideally situated to observe the IMF directly upstream of the Earth's magnetosphere and bow shock. The Imperial College magnetometer was operated for periods of up to 10 h during our experiments, and the

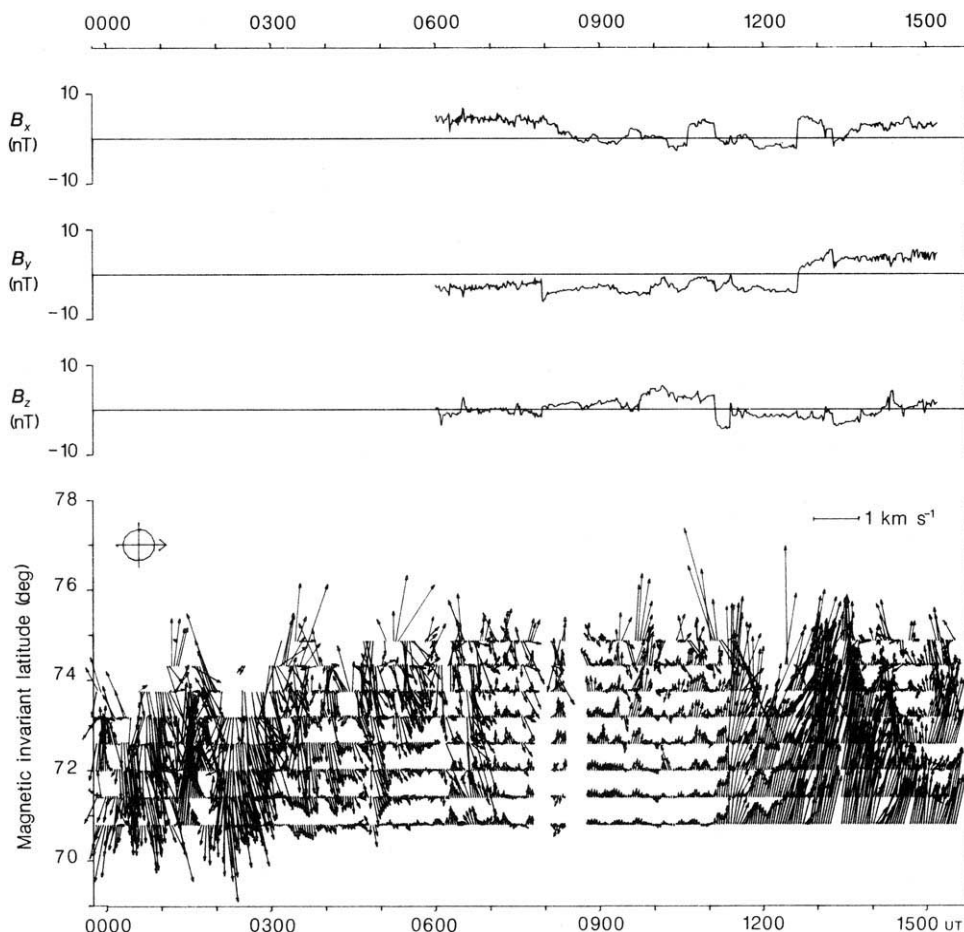
data were displayed in real time at the Rutherford Appleton Laboratory (RAL).

For the EISCAT observations we used the UK-Polar programme² to measure F-region plasma velocity at northwesterly ranges of 525(75)975 km from the UHF transmitter at Tromsø. At such ranges, the lines-of-sight to EISCAT's three receiving stations are so nearly parallel that it is necessary to use 'beam-swinging' to measure vector velocities^{3,4}. To reach high magnetic invariant latitudes Λ (71–75°) we pointed the Tromsø antenna at elevation 21.5°, and recorded data alternately at geographical azimuths 332° and 356° in a 5-min cycle. By combining the line-of-sight components measured in the two beam directions, we obtained two-dimensional velocity vectors, which were plotted at RAL shortly after each measurement. This procedure assumes that the plasma velocity does not vary greatly with longitude over distances of the order of 400 km, nor in about 5 min of time, and that the velocity parallel to the geomagnetic field ($\mathbf{B}$) is negligible. At high latitudes the plasma velocity is essentially the $\mathbf{E} \times \mathbf{B}$ drift due to a large-scale electric field $\mathbf{E}$, and we believe these assumptions do not seriously influence our conclusions.

The coordinated experiments ran for several hours on 29 September, 25 October and 27 October 1984. Here we discuss features of particular interest seen on 27 October, a day of fairly low geomagnetic activity ($\Sigma K_p = 16+$). Figure 1 shows the data sets covering the periods 0600–1510 UT for AMPTE-UKS and 0000–1530 UT for EISCAT. To minimize congestion of the vectors, we plot the EISCAT data in 'electric field' format, in which upwards on the diagram represents a magnetically northward electric field, corresponding to westward plasma drift velocity.

In high latitudes, the ionospheric plasma convection usually has a twin-cell pattern, consisting of a day-to-night flow across the polar cap, with return flows outside the polar cap on both dawn and dusk sides^{5–7}. Because EISCAT cannot normally

Fig. 1 Interplanetary magnetic field and ionospheric plasma drift velocity on 27 October 1984. Data from the Imperial College magnetometer on the AMPTE-UKS spacecraft are plotted as 1-min averages of the Cartesian components B_x , B_y , B_z in the GSM system (in which x points towards the Sun, and the x - z plane contains the geomagnetic dipole vector). Plasma velocity vectors measured by EISCAT are plotted with westward drift (corresponding to northward electric field) pointing upwards, northward drift to the right. In the EISCAT field of view, magnetic local time is approximately 2 h ahead of UT.



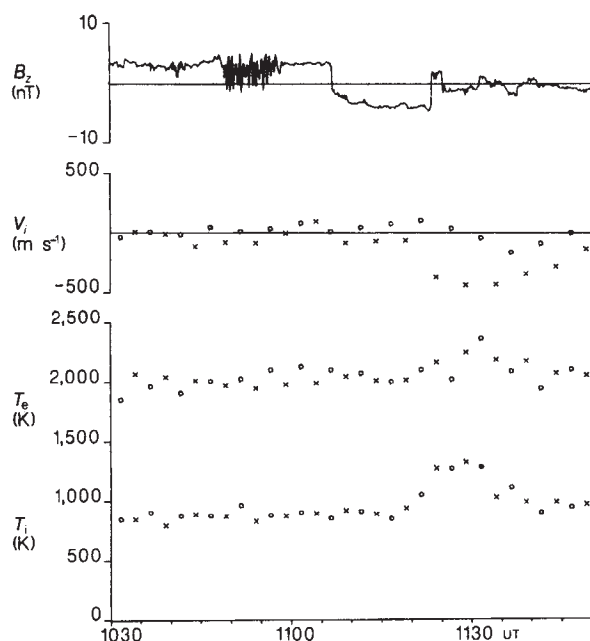


Fig. 2 Data for 1030–1150 UT on 27 October 1984, showing 5-s averages of the B_z component of the interplanetary magnetic field measured by AMPTE–UKS; and ionospheric plasma drift velocity, ion temperature and electron temperature measured by EISCAT at 750 km range, 311 km height, invariant latitude 72.6° (x, Geographical azimuth 332° ; o, azimuth 356°). The 40-s oscillations in B_z between 1049 and 1059 UT are probably 'upstream waves' associated with ions reflected from the bow shock¹⁴. According to a recent recalibration, the values of B_z shown in this figure should be reduced by 1 nT.

observe the polar cap by day, our measurements mostly show the return flows, eastward in the morning and westward in the evening. Depending on solar-geophysical conditions, the reversals between eastward and westward flow may occur some hours away from noon and midnight.

From 2350 UT we see predominantly fast eastward plasma flow ($1\text{--}2\text{ km s}^{-1}$), associated with the dawn side of the convection pattern. Between 0230 and 0310 UT this fast-flowing region recedes across the field of view, the poleward speed of the boundary being $\sim 150\text{ m s}^{-1}$ which could be produced by the Earth's rotation beneath a fixed flow pattern. There follow several hours of quieter flow ($0.3\text{--}1\text{ km s}^{-1}$), during which the morning east-to-west reversal of plasma velocity occurs at around 0600 UT. From 0600 the IMF northward component (B_z) fluctuates about zero, becoming positive after 0800 UT.

At 1107 UT the IMF suddenly turns southwards and B_z is mainly negative (0 to -5 nT) for the next 3 h. This change is followed at 1120 UT by a sudden acceleration of the ionospheric plasma flow which we judge to be simultaneous (to within $\pm 1\text{ min}$) over the whole radar field of view. The coherence of the data suggests that the effect is temporal, the plasma being accelerated in $<5\text{ min}$ to speeds of $\sim 1\text{ km s}^{-1}$ over a range of at least 3° in Λ . In principle, the rapid flow could have started earlier in the auroral oval (probably at $\Lambda \sim 76^\circ$ in quiet conditions⁸) and expanded across the field of view, but that seems unlikely. It would require an equatorward boundary motion of at least 3 km s^{-1} , 10 times what is typically observed⁹. Similar plasma accelerations have been observed by the Sondrestrom radar¹⁰, but with a coarser time resolution of $\sim 25\text{ min}$.

It is reasonable to suppose the southward IMF turning and the acceleration of ionospheric plasma are related, through processes that involve field-line reconnection at the dayside magnetopause. If so, roughly half the observed 13-min delay can be attributed to propagation from the satellite to the ground,

as follows: (1) Travel-time of IMF disturbance from AMPTE–UKS to the subsolar magnetopause: $4 \pm 1\text{ min}$; (2) propagation time of Alfvén waves from the subsolar magnetopause to the cusp ionosphere: $2 \pm 1\text{ min}$.

The other half could be due to three factors. First, the onset of reconnection might be delayed after the arrival of southward-pointing field at the magnetopause, although we do not know why that should happen. Second, several Alfvén transits may well be required to communicate to the ionosphere the full stress imposed by reconnection¹¹. Third, up to 3 min extra delay would occur if (as mentioned above, but considered unlikely) the rapid flow is initiated by an expansion of the auroral oval, instead of being transmitted through the ionosphere in $\sim 1\text{ s}$ at the Alfvén speed. We note that a similar unexplained delay of about 10 min has been found between changes of the IMF B_z and changes of the DP2 current system, as measured by ground-based magnetometers¹². DP2 is the magnetic disturbance associated with the twin-cell plasma convection pattern, which we measure in a more direct way using the radar.

The increase of plasma drift velocity is accompanied by a rise of 400 K in ion temperature between 1120 and 1130 UT (Fig. 2), which we attribute to ion-neutral frictional heating. On the timescale we are considering, the difference between ion and neutral gas temperatures, T_i and T_n , is given approximately by

$$(V_i - U)^2 = (3R/M)(T_i - T_n) \approx 1,300(T_i - T_n) \quad (1)$$

(see ref. 13), where V_i is the plasma drift velocity, U the neutral-air velocity, R the gas constant and M the mean molecular mass of the air (taken as 19 AMU). Assuming that initially $U \approx V_i \approx 200\text{ m s}^{-1}$ and that $T_i \approx T_n$, as seems reasonable, the observed temperature rise of 400 K requires $|V_i - U| \sim 700\text{ m s}^{-1}$, well within our observations. The ion temperature subsequently decreases, presumably because $(V_i - U)$ decays as the neutral air is accelerated by collisions with the drifting ions.

To summarize, we have observed a southward turning of the IMF, followed 13 min later by acceleration of ionospheric plasma. We believe this event demonstrates the connection between ionospheric plasma flows and the IMF direction, although the observed delay is some minutes more than expected on the basis of simple propagation from the satellite to the ground. The ionospheric temperature data support our interpretation. This event is the best isolated southward turning that we observed. We shall discuss elsewhere other, less clear-cut changes of IMF in our data, and we think that the ionospheric response may depend on factors such as local time. We stress the importance of making simultaneous IMF and ionospheric observations with good spatial and temporal resolution.

We thank the director and staff of EISCAT for help, Dr D. J. Southwood for the provision of AMPTE–UKS magnetometer data and Dr W. A. C. Mier-Jedrzejowicz for aid in magnetometer data processing. EISCAT is supported by the British SERC, French CNRS, West German MPG, Norwegian NAVF, Swedish NFR and Finnish SA. AMPTE is a collaborative project of NASA (USA), DFVLR and MPG of West Germany and SERC of the UK. P.R.S. is supported by an SERC grant.

Received 9 July; accepted 4 September 1985.

- Southwood, D. J. & Bryant, D. A. *Nature* **312**, 594 (1984).
- van Eyken, A. P., Rishbeth, H., Willis, D. M. & Cowley, S. W. H. *J. atmos. terr. Phys.* **46**, 635–641 (1984).
- Evans, J. V., Holt, J. M. & Wand, R. H. *J. geophys. Res.* **84**, 7059–7074 (1979).
- Foster, J. C., Doupinik, J. R. & Stiles, G. S. *J. geophys. Res.* **86**, 11357–11371 (1981).
- Heelis, R. A., Lowell, J. K. & Spiro, R. W. *J. geophys. Res.* **87**, 6339–6345 (1982).
- Cowley, S. W. H. in *High Latitude Space Plasma Physics* (eds Hultqvist, B. & Hagfors, T.) 225–249 (Plenum, New York, 1983).
- Jørgensen, T. S. *et al. Geophys. Res. Lett.* **11**, 887–890 (1984).
- Meng, C.-I. in *Dynamics of the Magnetosphere* (ed. Akasofu, S.-I.) 23–46 (Reidel, Dordrecht, 1980).
- Horwitz, J. L. & Akasofu, S.-I. *J. geophys. Res.* **82**, 2723–2734 (1977).
- Clauer, C. R. *et al. Geophys. Res. Lett.* **11**, 891–894 (1984).
- Willis, D. M. *Planet. Space Sci.* **18**, 749–769 (1970).
- Nishida, A. & Maezawa, K. *J. geophys. Res.* **76**, 2254–2264 (1971).
- Perraut, S., Brekke, A., Baron, M. & Hubert, D. *J. atmos. terr. Phys.* **46**, 531–543 (1984).
- Hoppe, M. M., Russell, C. T., Frank, L. A., Eastman, T. E. & Greenstadt, E. W. *J. geophys. Res.* **86**, 4471–4492 (1981).

Geomagnetic secular variation in Sicily and revised ages of historic lavas from Mount Etna

J. C. Tanguy, I. Bucur & J. F. C. Thompson

Laboratoire de Géomagnétisme, Université de Paris 6 et CNRS, 94107 Saint Maur des Fossés Cedex, France

The variation of geomagnetic field direction in Sicily during the past 700 yr has tentatively been determined using lavas of known date from Mount Etna¹. Additional palaeomagnetic studies on several hundred volcanic samples, combined with archaeomagnetic investigations carried out on Norman buildings, have improved the previous results and permit a reconstruction of the geomagnetic variation curve to about AD 1000. This curve agrees well with those obtained for other European countries²⁻⁶ and may be used as a reference for checking the ages attributed to archaeological structures as well as volcanic products in southern Italy during the past 1,000 yr. The present results cast serious doubts on the true ages of numerous historically dated lavas from Mount Etna, most of which are at least several centuries older than previously believed. The conclusions have implications for the succession of eruptions, effusion rates, magmatic evolution, and so on, and demonstrate the inconsistency of eruptive models based on historical records alone.

The Etnean lavas from major eruptions during the past 400 yr are well known, the most famous being the 1669 flow that buried about 15 villages and part of the town of Catania. Contrary to popular belief, however, older eruptions are far from being accurately described^{7,8}. Between AD 1600 and 1300, the products of only three eruptive periods can be identified with a reasonable degree of confidence: (1) the 1536 and 1537 flows on the south flank⁹⁻¹¹; (2) the 1408 flow north of Pedara village^{11,12}; and (3) the 1329 cone of Monte Rosso near Fleri (east flank), the name of which is cited in two early manuscripts found by Recupero¹². The palaeomagnetic direction given by welded scoriae from Monte Rosso is consistent with that of a lava flow erupted in 1301-02 on the island of Ischia (Gulf of Naples)¹³: because this eruption is the only one that occurred at Ischia in historical times, its products cannot be confused with others. For com-

parison with Etnean lavas, the magnetic inclination of the Ischia flow has been corrected to the latitude of Sicily.

The lavas mentioned above have been used to reconstruct the geomagnetic variation curve of field direction (Fig. 1), no reference being made to direct measurements as they only extend back to AD 1890 in Sicily. Other flows of supposedly known age, which are shown with quotation marks in Fig. 1, have also been extensively sampled: their peculiar behaviour will be discussed later.

Detailed results and discussion of the techniques of study (such as the removal of viscous remanent magnetization and alternating field demagnetizations) are reported in refs 1 and 7. Note that Etnean lavas are quite sensitive to parasitic magnetizations acquired during sawing or coring. Therefore, we used only 'big' samples (0.5-2 kg) oriented by the plaster method^{1,2}, with azimuth reference from the Sun shadow. This technique also has the advantage of permitting a much better precision during sampling and measurements. Table 1 summarizes the results obtained after magnetic cleaning from lavas of both reliable and dubious dates. The semi-angle of the 95% confidence cone of the average flow directions ranges from 0.71 to 1.76°, thus giving a precision similar to that of usual archaeomagnetic data.

Figure 1 shows that the points corresponding to all the recent flows (AD 1910 to 1610) reproduce the curve of geomagnetic secular variation known from direct measurements in Paris and London^{2,3}, allowing for difference in latitude. By going further back, there is an increasing number of flows (indicated between quotation marks) whose palaeomagnetic directions are inconsistent with that of the geomagnetic field expected at the respective dates. The eruptions to which these lavas were ascribed are poorly located by the historical data. Sometimes, re-examination of early documents enables us to preclude the dates attributed—later—in the modern literature. For example, the "1595" flow, which corresponds to two petrologically distinct flows⁷—"Gallo Bianco East" and "Gallo Bianco West"—in Fig. 1, is undated in documents published before the middle of the nineteenth century. Furthermore, contemporaneous witnesses pointed out that Etna had been inactive for at least 30 yr before 1603 (refs 11, 14). The "1595" date is only quoted by Sartorius in his Atlas of Etna¹⁵, although this author does not mention any eruption for this particular year¹⁶. It would appear that "1595" is a printing error. A full discussion of the problems related to spurious dates attributed to Etnean lavas is given elsewhere^{7,8}.

Fig. 1 Geomagnetic variation in Sicily recorded in historically dated Etnean lavas. Ovals indicate the 95% confidence intervals. The dashed lines express uncertainties on the trend of the curve, 1284(?) and "1408" (Trecastagni) being possibly of prehistoric age. Dates in quotation marks are erroneous, although the morphology of the flows indicates a very recent age (see text). Note the general agreement with the archaeomagnetic curve obtained in France (upper left, from refs 2, 17).

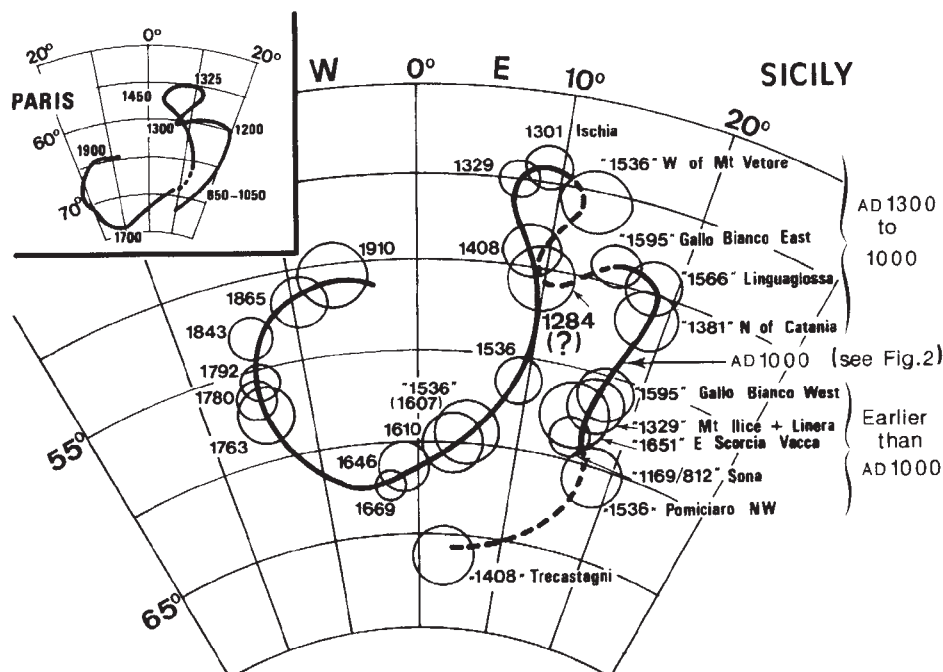


Table 1 Palaeomagnetic directions of lava flows or cones of known (a) or presumed (b) ages

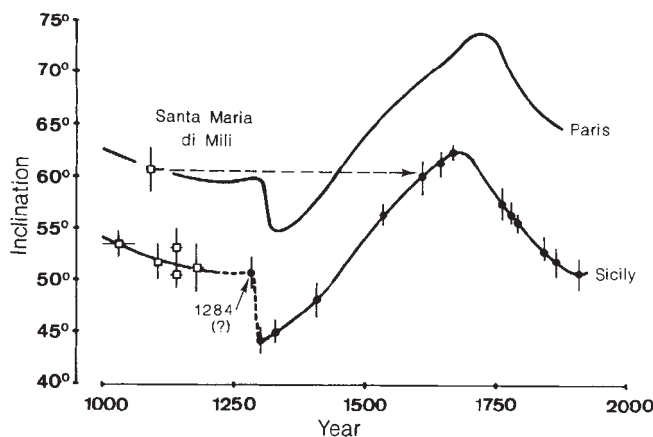
	<i>N</i>	$\bar{I}$ (deg)	$\bar{D}$ (deg)	α_{95}	<i>k</i>
a, Volcanic units of known age					
1669 I	7	61.7	-3.0	0.97	2,963
1669 II + III	7	62.3	-3.6	2.29	535
1669 IV	5	62.7	-3.0	1.62	1,503
1669 V	6	62.6	-2.7	1.35	1,793
1669 VI	5	63.8	-3.3	1.28	2,395
1669 (mean)	30	62.5	-3.3	0.71	1,310
1910	11	50.7	-6.9	1.53	760
1865	18	51.9	-9.5	1.41	583
1843	12	53.2	-14.2	1.12	1,552
1792	23	55.8	-14.2	1.01	837
1780	32	56.7	-14.9	1.04	582
1763	6	57.6	-14.4	1.51	1,437
1646	8	61.3	-1.5	1.25	1,564
1610	10	60.0	3.5	1.61	750
1536, Piano San Leo	21	56.5	8.7	1.04	855
1408, Pedara	12	48.1	7.9	1.67	703
1329, Mt Rosso, Fleri	10	45.1	7.1	1.04	1,824
1301-1302, Ischia	24	44.2	9.0	1.20	566
1284(?) Petrulli-Monacella	18	50.7	9.7	1.52	468
b, Volcanic units of presumed age					
"1651", Scorciovacca	12	57.5	14.7	1.75	581
"1595 I", Gallo Bianco West	16	56.0	17.0	1.44	591
"1595 II", Gallo Bianco East	10	48.9	15.1	1.20	1,351
"1566", Linguaglossa	19	49.7	17.7	1.40	525
"1536 I", West of Mt Vetore	4	45.7	12.6	1.76	1,583
"1536 II", Lower Pomiciaro (north-west)	9	60.7	18.3	1.58	874
"1536 III", Upper Pomiciaro	7	59.4	4.8	1.67	1,172
"1408", Treccastagni North	10	66.1	3.0	1.52	1,060
"1381", Gravina di Catania	12	51.1	18.6	1.41	815
"1329", Mt Illice, Linera	15	56.7	16.4	1.44	731
"1169", Aci Castello	11	57.3	8.2	1.24	1,165
"1169 or 812", Sona (south-west)	8	58.7	14.8	1.04	2,578

N is the number of samples for each eruptive unit (for the 1669 flows, the detail of each site of sampling is shown); α_{95} and *k* are the Fisher 95% confidence limits and precision parameters, respectively. $\bar{I}$ is the mean inclination and $\bar{D}$ the mean declination.

Thus, most of the lavas believed to have erupted between AD 1651 and 1284 are actually older than this period. The 1284(?) flow itself is problematic; both petrographic and palaeomagnetic evidence⁷ show that it consists of several flows of undoubtedly different ages in the area indicated as a single "1284" lava by Sartorius¹⁵. One of these flows has a palaeodirection that is not inconsistent with the trend of the geomagnetic variation expected for this late thirteenth-century period (compare with Paris and Oxford, Fig. 1 and refs 3, 17). However, its chemical composition and mineralogy are close to those of some prehistoric lavas, and its morphology is less preserved than for any of the historical flows. Thus, there is no objective argument to ascertain whether it could represent the 'true' 1284 lava.

The other flows of presumed age, in contrast, have a well-preserved morphology and petrological characteristics similar to those of well-dated lavas. They probably belong to unrecorded eruptions that occurred during the medieval period, an assumption which is consistent with the results of archaeomagnetic records obtained in other parts of Europe^{2-6,17}. These Etnan flows can be used, therefore, to reconstruct the trend of the geomagnetic curve (right-hand side of Fig. 1), although its variation as a function of time cannot be determined, owing to the lack of historical records.

For dating these Middle Age flows of Etna, at least approximately, a preliminary archaeomagnetic study has been carried out on bricks from the walls of six Norman buildings in Sicily (Table 2). Alternating-field cleaning and thermal demagnetization techniques² were applied to the brick samples; the result was an improvement of the 95% confidence intervals, while the mean inclination remained virtually unchanged. The curve of

**Fig. 2** Variation of geomagnetic inclination in Sicily during the past 1,000 yr, from lavas (solid symbols) and bricks (open symbols). The variation of inclination at Paris^{2,17} is shown for comparison.**Table 2** Geomagnetic inclination recorded in bricks from Norman buildings in Sicily

Site	Date (AD)	No. of samples	$\bar{I}$ (deg)	Treatment	Observations
Santi Pietro e Paolo di Agro	≈ 1000-1060	43	53.5	—	Bricks (church constructed before the Norman period)
		41	53.0 ± 1.4	15 mT	
Santa Maria di Mili	"1092"	25	60.7 ± 2.1	—	Bricks (foundation of church, 1092; possible restoration ~1542)
Santissimo Salvatore di San Marco d'Alunzio	1105	31	51.8 ± 1.7	—	Bricks (monk Gregorio's testament: foundation of church 1105)
Palazzo Reale (Palermo)	1130-1150	47	50.5 ± 1.3	—	Bricks (Pisan Tower windows)
Cefalù	1132-1148	17	53.5 ± 2.1	—	Bricks from the south nave of the church
		17	53.2 ± 1.8	150 °C	
Monreale (Palermo)	1172-1185	17	51.8 ± 2.6	—	Bricks (ornamental material from terrace)
		17	51.2 ± 2.3	140 °C	

geomagnetic inclination obtained (Fig. 2) is very similar to that observed in France (refs 2, 17 and I.B., unpublished data), with the difference of latitude between the two countries taken into account. An unexpected archaeological result concerns the Santa Maria di Mili church: although built in AD 1092, the church has clearly been restored during the sixteenth century, as revealed by the strong magnetic inclination carried by all the bricks sampled.

This new inclination curve for Sicily shows that the supposed Etna flows of "1595" (Gallo Bianco East), "1566" (Linguaglossa) and "1381" (north of Catania) have palaeomagnetic inclinations consistent with eruptions occurring at ~AD 1050-1250. Other flows (lavas from Gallo Bianco West, Mount Illice, Scorciovacca, Sona, Pomiciaro north-west and Treccastagni north) are significantly older, probably between AD 700 and 1000, judging

from the strong increase of inclination recorded at this epoch in Europe^{2-6,17}.

We hope to improve our knowledge of the geomagnetic variation in Sicily by investigating the archaeological structures dated from the Greek civilization to the Roman and Byzantine periods. The resultant curve could be used to check the true age of numerous eruptive products from South Italian volcanoes, as well as archaeological structures of uncertain date.

We thank M. J. Aitken and M. Prévot for advice and criticism. This study is a contribution of the French program of Volcanology (PIRPSEV, CNRS-INAG), in the context of a French/Italian cooperation (GNV, CNR).

Received 10 June; accepted 23 September 1985.

1. Tanguy, J. C. *Archaeometry* 12, 115-128 (1970).
2. Thellier, E. *Phys. Earth planet. Inter.* 24, 89-132 (1981).
3. Aitken, M. J. *Phil. Trans. R. Soc. A239*, 77-88 (1970).
4. Kovacheva, M. *Geophys. J. R. astr. Soc.* 61, 57-64 (1980).
5. Turner, G. M. & Thompson, R. *Earth planet. Sci. Lett.* 42, 412-426 (1979).
6. Creer, K. M., Readman, P. W. & Papamarinopoulos, S. *Geophys. J. R. astr. Soc.* 66, 193-219 (1981).
7. Tanguy, J. C. Thesis, Univ. Paris 6 (1980).
8. Tanguy, J. C. *Bull. Volcan.* 44, 585-640 (1981).
9. Fazellus, T. *De rebus siculis*, IIII, (Maida, Panormi, 1558).
10. Philoteus, A. *Aetnae topographia atque ejus incendiorum historia* (Muschi, Venice, 1590).
11. Carrera, P. *Il Mongibello in tre libri* (Rossi, Catania, 1636).
12. Recupero, G. *Storia Naturale e Generale dell'Etna* Vol. 2 (Stampa Regia Università, Catania, 1815).
13. Tanguy, J. C. & Pozzi, J. P. *C.r. hebdom. Séanc. Acad. Sci., Paris* 274D, 352-354 (1972).
14. Monaco, F. *Cataclysmus Aetneus sive inundatio ignea Aetnae Montis anni 1669*, 10 (Hertz, Venice, 1669).
15. Sartorius von Waltershausen, W. *Atlas de l'Etna* (Schropp, Berlin, 1845; Vandenhoek & Ruprecht, Goettingue, 1848; Geographisches Institut, Weimar 1859).
16. Sartorius von Waltershausen, W. *Der Aetna* Vol. 1 (W. Engelman, Leipzig, 1880).
17. Bucur, I. *Proc. 24th int. Archaeometry Symp.* (Smithsonian Institute, in the press).

A dynamic model of the curvature of the Mariana Trench

Albert T. Hsui* & Sarah Youngquist*†

* Department of Geology, and † Physics Department, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Oceanic trenches are prominent features of the Earth's surface. They represent the convergent plate boundaries and generally possess convex curvatures towards the subducting plates. Based on this observation, many static geometric models have been proposed for the origin of trench curvatures¹⁻⁶. Unfortunately, many of these models have limited applicability⁷. Alternatively, a dynamic model has been proposed which suggests that trench curvatures are evolved from collisions between aseismic ridges and the trench axis⁸. We have developed a flow model to quantitatively evaluate this dynamic process for the Mariana Trench system. The model is able to yield an excellent fit to the present trench curvature and determine past rotations for back-arc volcanic islands such as Saipan and Guam with surprising accuracy when compared with palaeomagnetic observations.

The Mariana Trench represents the convergent boundary between the East Philippine plate and the subducting Pacific plate. It is bounded to the north by the Marcus-Necker Ridge and to the south by the Caroline Ridge (Fig. 1). Both ridges are classified as aseismic ridges because of their elevated bathymetry and their lack of large seismic activities^{9,10}. These ridges are more buoyant relative to their adjacent oceanic sea floor due to either thermal (for example, higher temperature) or chemical (for example, thicker crust) effects. Because of their stronger buoyant character, they tend to resist subduction⁵. As a result, these aseismic ridges will continue their westward migration while other parts of the ocean floor begin their descent into the mantle. Thus, an initially straight trench will evolve into a convex geometry towards the subducting plate. In this scenario, the overriding plate is assumed to be stationary while the subducting plate is converging. However, the same problem can be formulated in which the subducting plate is chosen to be stationary

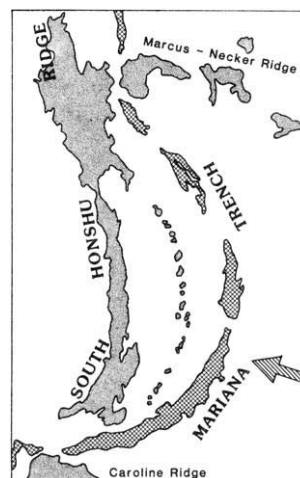


Fig. 1 Tectonic setting of the Mariana Trench.

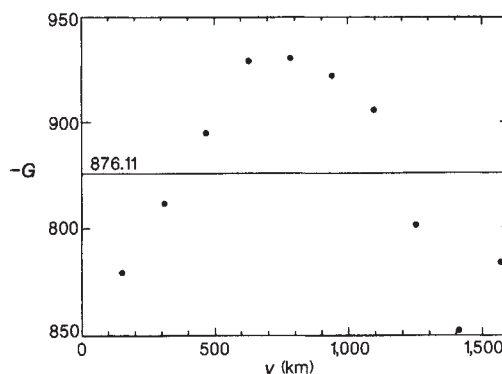


Fig. 2 Observed values of G as a function of y where y is defined in Fig. 3. The solid line represents the mean value for G .

while the overriding plate is flowing over it. As far as the plate boundary geometry is concerned, these two reference coordinates are interchangeable. However, the latter view provides some mathematical simplicities in formulating the problem.

Consider a newtonian viscous fluid with a linear front moving over two stationary point obstacles. The obstacles represent the leading edges of the aseismic ridges as shown in Fig. 1. Since viscosity of geological materials is quite high especially in surface conditions, surface deformation will obey Stokes' flow where inertia forces can be neglected. The governing equation is then given by

$$0 = -\frac{1}{\rho} \frac{dP}{dx} + \nu \frac{d^2 u}{dy^2} \quad (1)$$

where dP/dx is the horizontal pressure gradient arising from stress imbalance within the overriding plate, possibly due to the unboundedness of the trench wall, and it is assumed to be constant; ν is the kinematic viscosity and u is the flow velocity. As the obstacles are stationary, boundary conditions require vanishing velocity at $y = 0$ and $y = L$. L is the separation between the obstacles. Consequently, equation (1) can be integrated twice to obtain a velocity profile which in turn can be integrated again with respect to time to yield a curvature profile as given below, where μ is the viscosity of the fluid layer:

$$x = \frac{1}{2\mu} \frac{dP}{dx} (y^2 - Ly) t \quad (2)$$

This equation describes the evolution of a trench curvature according to the model of Vogt *et al.*⁸.

As there are some parameters in the model which are not very well constrained, we shall first attempt to use observed data to determine the acceptable range of values for these parameters. Later, we shall demonstrate that values thus determined are able to fit another set of geophysical data. By doing so, the validity of our model can be established.

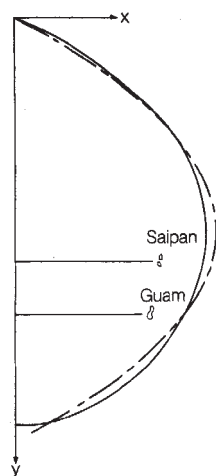


Fig. 3 Curvature of the Mariana Trench. The solid line is the observed curvature while the dashed line is the theoretical curvature based on the mean value of G as determined in Fig. 2.

Evolution of the Mariana Trench has been studied by many investigators^{11,13}. The collision between the aseismic ridges and the proto-Mariana trench is set at approximately the middle Miocene time. While some constraints can be established for the timescale, very little can be said about the newtonian viscosity of the surface layer or the pressure gradient across a trench, which are the other two parameters intrinsic to the model. Fortunately, these three parameters can be combined to form a single parameter group [that is, let $1/G = (1/2\mu) (dP/dx)t$]. We shall test if the observed Mariana Trench curvature is able to yield a constant G . Taking the separation between the two aseismic ridges to be $\sim 1,650$ km, the observed values for G are given in Fig. 2. The mean value for this data set is -876.11 km. Based on this mean value, a theoretically predicted trench curvature can be constructed according to equation (2). Comparison of the predicted and observed profiles is given in Fig. 3. Agreement between the model and observation is quite good.

To demonstrate the validity of this model, we shall next examine the relationship between this model and some palaeomagnetic studies behind the Mariana Trench. Application of palaeomagnetic data to study plate rotation has been carried out by many investigators^{11,14-17}. Palaeomagnetic declinations of volcanic rocks of different ages can be measured at the same location. The discordance of palaeomagnetic direction will yield information on the sense and magnitude of rotation for the sampled localities. McCabe¹¹ has compiled some palaeomagnetic data for the southern Mariana Islands. Our model predictions will be compared with these data.

Investigation of a fluid particle rotation within a flow such as the one described previously, requires the theory of finite deformation¹⁸. Geological applications of finite deformation theory have been carried out by Elliot¹⁹ and McKenzie²⁰. In our case, as the flow field is very similar to a steady simple parallel flow, the rate of finite rotation of any fluid particle within the flow can be described by a linear combination of shear strain and fluid vorticity¹⁹, where e is shear strain and ω is vorticity.

$$\theta = e + \omega = \frac{du}{dy} = \frac{1}{2\mu} \frac{dP}{dx} (2y - L) \quad (3)$$

Consequently, the finite angle of rotation can be written simply as the time integration of equation (3).

$$\theta = \frac{1}{2\mu} \frac{dP}{dx} (2y - L)t + \theta_0 \quad (4)$$

where θ_0 is the initial angle with respect to the magnetic north prevalent at the time of the collision. Equation (4) can also be derived using a more rigorous approach as described in refs 18, 20. Based on equation (4), finite rotation of the volcanic islands

can be calculated. Note that the time variable t here is not the same as that in equation (2) for trench curvature determination; it is the time that is required for the island to travel from the point when it crossed the line connecting the two obstacles to its present position. Given the flow velocity and the distance, the time duration can be estimated. As a result, rotation of volcanic islands as a function of time can be obtained, and is plotted in Fig. 4; also given are the palaeomagnetic rotation data for Guam and Saipan Islands as compiled by McCabe¹¹. Agreement is surprisingly good.

As indicated in Fig. 3, a parabolic profile is adequate to fit the Mariana Trench. Based on a mean value of -876.11 km for the grouped parameter G , a fit between the theory and the observation can be obtained with a standard deviation of $<5\%$. A closer examination of the comparison indicates that the mismatch is not randomly distributed. The theory over-predicts movement near the centre part of the trench while underestimations occur near the aseismic ridges. One possible reason for this discrepancy may be the existence of a small component of non-newtonian rheology within the overriding plate, since power creep laws will tend to flatten a parabolic profile as evidenced in glacial flow²¹ and in other non-newtonian channel flows²². The non-newtonian contribution appears to be small, however, since the match by a newtonian theory is quite sufficient. Alternatively, the bathymetric trench axis may not represent the actual convergent boundary due to the accumulation of sediments in the accretionary prism²³. Unfortunately, the actual boundary geometry can only be mapped with dense coverage of seismic reflection experiments.

Based on tectonic reconstruction studies, the initial collision between the aseismic ridges and the Mariana Trench has been estimated to have occurred at about the middle Miocene^{11,13}. We thus assign a 15 Myr numerical value for computational purposes. It follows, therefore, that the mean value for $(1/2\mu)/(dP/dx) = 2.4 \times 10^{-21} (\text{m s})^{-1}$. Using this best-fit value, we are able to calculate the velocity of the two volcanic islands (Guam and Saipan), where palaeomagnetic rotation data are available for our model comparison. Table 1 gives the values for velocity and elapsed time for the two volcanic islands. Using

Table 1 Parameter values used for palaeomagnetic rotation calculation

	Guam	Saipan
θ_0 (deg)	-19°	-19°
L (km)	1,650	1,650
Time since collision (Myr)	15	15
x (km)	515.63	556.88
y (km)	1,155.00	948.75
Velocity (km Myr ⁻¹)	43.31	50.88
Time since rotation started (Myr)	11.91	10.95
θ (deg)	-53.26°	-31°

these values, the rotational history for Guam and Saipan can be calculated. Excellent agreement between our model predictions and the observations is clearly demonstrated in Fig. 4. In addition, our model is able to confirm the fact that rotation at Guam is more intense than that at Saipan. It is simply because Guam is closer to the Caroline Ridge. As a result, more vorticity and shearing are expected. Saipan, on the other hand, is closer to the midpoint between the two aseismic ridges where flow is more uniform and with vanishing shear and vorticity. Note that the amount of rotation as determined by our model depends only on the time interval between the present and the time at which the volcanic islands flowed past the leading edges of the aseismic ridges, if the rocks were formed earlier. If rock samples were formed after that time, rotation is calculated based on the time interval between the present and the time at which the rocks were solidified. Consequently, rotation of younger rocks can also be calculated. Unfortunately, data for younger volcanic rocks are not sufficient for a meaningful comparison. Of course, maximum rotation behind the migrating trench will occur in the mid-Miocene volcanics. Rocks with ages older than mid-

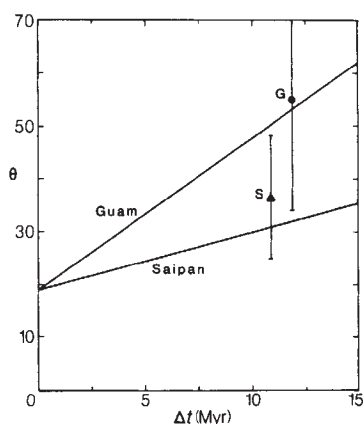


Fig. 4 Palaeomagnetic rotation as a function of time for Guam and Saipan islands. Observed rotations (●, ▲) as reported in ref. 11 are also plotted.

Miocene are not expected to show any larger rotation, since the trench is hypothesized to have moved uniformly before the collision with the two aseismic ridges. Because of the excellent agreement between our model and the palaeomagnetic data, the proposed dynamic model for the curvature of the Mariana Trench is believed to be valid. The model is especially attractive because of its ability to explain the differential rotation between Saipan and Guam.

The remaining issue is whether the value of 2.40×10^{-21} chosen for $(1/2\mu)/dP/dx$ is reasonable. To address this question, it is necessary to know the viscosity of the surface layer and the pressure gradient within the overriding plate; neither parameter is very well understood, thus one can only examine the range of possibilities. If the pressure gradient is of the order of a few hundred bars over 1 km, then the viscosity will be about 10^{26} . A pressure difference of a few hundred bars is believed to be reasonable as this is about the strength of surface rocks. Is the required viscosity too low? Although very few studies have been carried out to estimate the viscosity of the surface layer, some laboratory observations have been made in an attempt to provide answers to this question²⁴. Based on Ito's experiment on creep deformation of a granite slab over a 10-yr period²⁴, he found that the deformation of his granite slab can practically be described by a newtonian rheology with a viscosity of $\sim 10^{23}$ Pa s. Thus, our implied value for surface rock rheology is not unreasonable in view of the values derived from Ito's experiments.

We thank Dr Stephen Marshak for many helpful discussions, Dr Peter R. Vogt for his critical review. S.Y. is supported by a graduate fellowship from the AT & T Bell Laboratories.

Received 3 July; accepted 30 September 1985.

1. Frank, F. G. *Nature* **220**, 363 (1968).
2. Van der Held, E. F. M. *Proc. K. ned. Akad. Wet.* **B72**, 256-286 (1969).
3. Strobach, K. Z. *Geophys.* **39**, 819-831 (1983).
4. De Fazio, T. L. *Tectonophysics* **23**, 149-154 (1974).
5. Laravie, J. A. *Geology* **3**, 484-486 (1975).
6. Peltzer, G. *Earth planet. Sci. Lett.* **55**, 463-472 (1981).
7. Tovish, A. & Schubert, G. *Geophys. Res. Lett.* **5**, 329-332 (1978).
8. Vogt, P. R., Lowrie, A., Bracey, D. R. & Hey, R. N. *Spec. Pap. Geol. Surv. Am.* (1976).
9. Fryer, P. & Smoot, N. C. *Mar. Geol.* **64**, 77-90 (1985).
10. Keating, B., Matney, D., Naughton, J., Epp, D. & Helsley, C. E. *EOS* **62**, 381-382 (1981).
11. McCabe, R. *Tectonics* **3**, 409-428 (1984).
12. McCabe, R. & Uyeda, S. *Geophys. Monogr. Ser.* **27**, 281-293 (1983).
13. Uyeda, A. & Ben-Avraham, Z. *Nature phys. Sci.* **240**, 176-178 (1972).
14. Irving, E. *Can. J. Earth Sci.* **16**, 669-694 (1979).
15. Beck, M. E. *Am. J. Sci.* **276**, 694-712 (1976).
16. Beck, M. E. *J. geophys. Res.* **85**, 7115-7131 (1980).
17. McKenzie, D. & Jackson, J. *Earth planet. Sci. Lett.* **65**, 182-202 (1983).
18. Malvern, L. E. *Introduction to the Mechanics of a Continuous Medium* (Prentice-Hall, New Jersey, 1969).
19. Elliott, D. *Bull. geol. Soc. Am.* **83**, 2621 (1972).
20. McKenzie, D. *Geophys. J. R. astr. Soc.* **58**, 689-715 (1979).
21. Hughes, T. *Rev. Geophys. Space Phys.* **15**, 1-46 (1977).
22. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
23. Karig, D. E. *A. Rev. Earth planet. Sci.* **2**, 51-75 (1974).
24. Ito, Hidebumi. *Tectonophysics* **52**, 629-641 (1979).

Palaeobiological and sedimentological implications of fossil concentrations

Susan M. Kidwell

Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, Illinois 60637, USA

Concentrations of fossil hardparts are common features of the stratigraphical record and are preferred collecting sites for most palaeontological data. Nonetheless, most investigations into the nature of the fossil record have analysed the biasing effects of selective hardpart transport and destruction^{1,2} rather than the consequences of the concentration process itself. Genetic classification is discouraged by the diverse origins of skeletal accumulations, which range from predator gastric residues to shelly shoals and biostratigraphically condensed deposits; concentrations can thus form over time intervals of a few minutes to hundreds of thousands of years. I show here that the close association of shell beds with stratigraphical discontinuities in Miocene shallow marine deposits of Maryland^{3,4} provides the basis of a model of skeletal accumulation cast entirely in terms of changes in net sedimentation. This simple sedimentological model is a surprisingly powerful predictor of post-mortem bias and ecological composition of fossil assemblages, suggesting that fossil-rich and fossil-poor strata are qualitatively different, both as repositories of palaeontological information and as settings for biotic interactions. Moreover, the apparent primary importance of rates of sedimentation in skeletal accumulation—despite emphasis usually placed on rates of hardpart input—suggests a new approach to inferring the detailed dynamics of sediment deposition and erosion in the formation of stratigraphical sequences.

Description of the upper and lower contacts of shell (or bone) beds as sharp or gradational has genetic significance because bed contacts not only describe the physical relation of the shell bed to adjacent beds, but also relate the process of hardpart concentration to processes responsible for the accumulation of the surrounding, less fossiliferous sediment. Except for contacts created by bioturbation or diagenesis alone, sharp contacts indicate disjunct shifts in sedimentation and usually record an

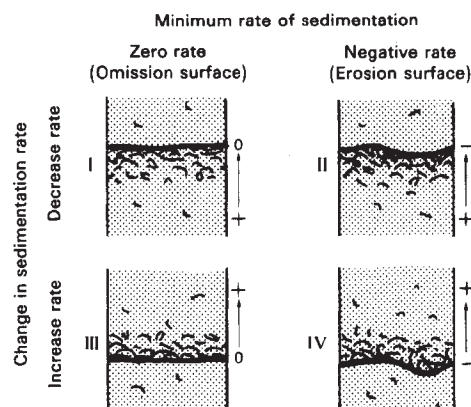


Fig. 1 Field classification of shell beds on the basis of stratigraphical contacts. Type I shell beds grade from less fossiliferous sediment and terminate in a sharp omission surface; type II shell beds also grade upwards with increasing shell-packing density but terminate along an erosion surface; type III shell beds rest on an omission surface and grade upwards into less fossiliferous sediment; and type IV concentrations have erosional basal contacts and grade upwards. Concentrations associated with such surfaces must arise in the context of some change in net sedimentation, since discontinuities indicate conditions of zero (omission) or negative (erosion) net sedimentation, and sediment between discontinuities records positive net sedimentation (deposition).

Table 1 Taphonomic and palaeobiological predictions for four simple patterns of fossil accumulation as modelled in Fig. 1

Shell bed type	Modelled change in sedimentation	Hardpart residence time on seafloor	Hardpart abundance on seafloor	Physical abrasion & fragmentation	Severity of post-mortem bias				Community composition			Species morphology		
					Bio-erosion	Dissolution	Transport addition or removal	Time-averaging	Deep or mobile infauna	Encrusting & boring taxa	Shell gravel taxa	Shell gravel morphs	Variance owing to: Time-averaging	Selective destruction
I	Decrease from + to 0	Increase	Increase	Increase	Increase	Increase	+/- Increase	Increase	Decrease	Increase	Increase	Increase	Increase	+/- Decrease
II	Decrease from + to -	Increase	Increase	Strong increase	Increase	Increase	Increase	Strong increase	Decrease	Increase	Increase	Increase	Strong increase	Decrease
III	Increase from 0 to +	Decrease	Decrease	Decrease	Decrease	Decrease	+/- Decrease	Decrease	Increase	Decrease	Decrease	Decrease	Decrease	+/- Increase
IV	Increase from - to +	Decrease	Decrease	Strong decrease	Decrease	Decrease	Decrease	Strong decrease	Increase	Decrease	Decrease	Decrease	Strong decrease	Increase

Hardpart abundance and hardpart residence time at or near the sea floor are treated as simple functions of rate of sedimentation (hardpart input is assumed to be constant); fossil assemblages sampled from different horizons within any particular shell bed will be characterized by different degrees of post-mortem bias and patterns of biotic response.

episode of erosion or omission (non-deposition)⁵⁻⁷. Hardpart concentrations associated with sharp discontinuities thus must have formed in the context of some change in net sedimentation. These concentrations can be divided into four basic types (Fig. 1), depending on whether hardparts lie on top of or directly underneath the discontinuity surface (as in the "Sohlbanke" and "Dachbanke" of Brinkmann⁸), and whether the surface exhibits scour and truncation features of erosion (negative sedimentation rate) or merely omission (zero net sedimentation rate). As an initial simplifying assumption, hardpart input is assumed to be constant throughout the accumulation period; the model is later shown to hold even after relaxation of this unrealistic assumption.

Type I and type II shell beds (Fig. 1) indicate hardpart accumulation during a slowing down in net sedimentation from high initial rates which exceed hardpart input. Hardpart packing density increases upwards as sedimentation approaches a zero or negative net rate. Type III and type IV beds record an acceleration in sediment accumulation from initially very low rates that foster the concentration of hardparts towards higher positive rates which exceed the background rate of hardpart input. These dynamics of sedimentation have immediate implications for assemblages collected from both fossil-rich and fossil-poor strata.

As rates of sedimentation decrease during the formation of types I and II shell beds, individual hardparts are exposed at the depositional interface for increasingly long periods of time. Assemblages collected from successively higher horizons within the shell bed should thus exhibit a higher frequency and severity of shell damage by abrasion, fragmentation and bio-erosion (Table 1). This trend would be reversed in types III and IV beds. In addition to suffering more pronounced bias owing to selective destruction of hardparts, assemblages that accumulate during the phase of lowest net sedimentation are also more likely to be time-averaged, that is, composite records of a series of populations that occupied the site at different points in time. Time-averaging should be more severe in types II and IV beds than in types I and III beds because of vertical mixing of faunas during erosional reworking. Types II and IV beds are also most likely to be biased by the addition and removal of hardparts through selective transport.

By controlling rates of hardpart burial, net sedimentation also determines the abundance of hardparts in the substratum and thereby influences the physical characteristics of benthic habitats. Ecologically, type I shell beds record the transformation of an initial soft-bottom habitat into an increasingly shelly and thus coarser-textured, firmer and topographically more complex substratum. This change in the physical environment facilitates colonization by borers and encrusters of dead hardparts and by free-living and attached epifauna which prefer or require stable substrata⁹. At the same time, the *in situ* development of a shell gravel inhibits mobile and sedentary burrowers which occupied

the original soft sediment habitat⁹. Types I and II beds should thus exhibit an upward increase in the abundance and diversity of epifaunal species and a concomitant decrease in the numbers of mobile and of especially large or deep-burrowing infauna (Table 1). Assemblages from the most densely fossiliferous parts of the beds will be characterized by ecologically mixed assemblages produced by later shell-gravel taxa occupying the same volume of sediment as earlier soft-bottom taxa; types III and IV beds should exhibit the opposite trend. These shifts in faunal composition should arise regardless of whether the dead hardparts that bring about the change in benthic habitat were produced *in situ* or delivered from allochthonous sources^{9,10}.

Another expected biotic response to variation in shell-packing density is a shift in average shell shape of morphologically plastic taxa—most notably epifauna—which can occupy both soft and shelly bottoms¹⁰. Shell gravel morphs should occur most frequently in assemblages from the most densely packed part of a shell bed (Table 1). Apparent morphometric variance within sampled fossil populations of species should also be greatest there, owing to the time-averaging of the directional shift from soft-bottom and shell gravel morphs and of random non-directional fluctuations in morphology. In types III and IV shell beds, the expected burst of variation followed by a gradual dwindling of variance mimics patterns predicted by some theories of speciation^{11,12}.

The robustness of the sedimentological model is evaluated by relaxing the assumptions, using a series of hypothetical histories of sedimentation and hardpart input (Fig. 2). When hardpart input is allowed to vary (Fig. 2b-e), classification and interpretation of shell beds using the bed contact criterion yields a correct interpretation of sedimentary dynamics in all but one situation: when peaks in hardpart input coincide precisely with maxima in sedimentation (Fig. 2d). This relationship is biologically improbable, because most shelly benthos avoid colonizing settings with high rates of sedimentation. Coincident maxima can occur when hardparts are hydraulically equivalent to enclosing sediments, such as macroinvertebrate shell debris in turbidity current or other high-energy conditions, and microfossil tests in deep-sea environments. These pitfalls can be diagnosed in the field by visual comparison of the hardpart and sediment grain sizes. Figure 2e, f illustrates the results of more complicated patterns of change in sedimentation and hardpart input and more closely mimics real fossiliferous sequences. The robustness of the model for these, as well as for endmember conditions, suggests that the model can be applied successfully to the spectrum of possible combinations of rate changes.

Changes in net sedimentation inferred from bed contacts provide working hypotheses that can be tested by their palaeontological predictions (Table 1) as well as by independent evidence for reduced net sedimentation within the shell bed (glauconite concentration, winnowed sedimentary matrix, physically amalgamated beds). In the Maryland Miocene sequence,

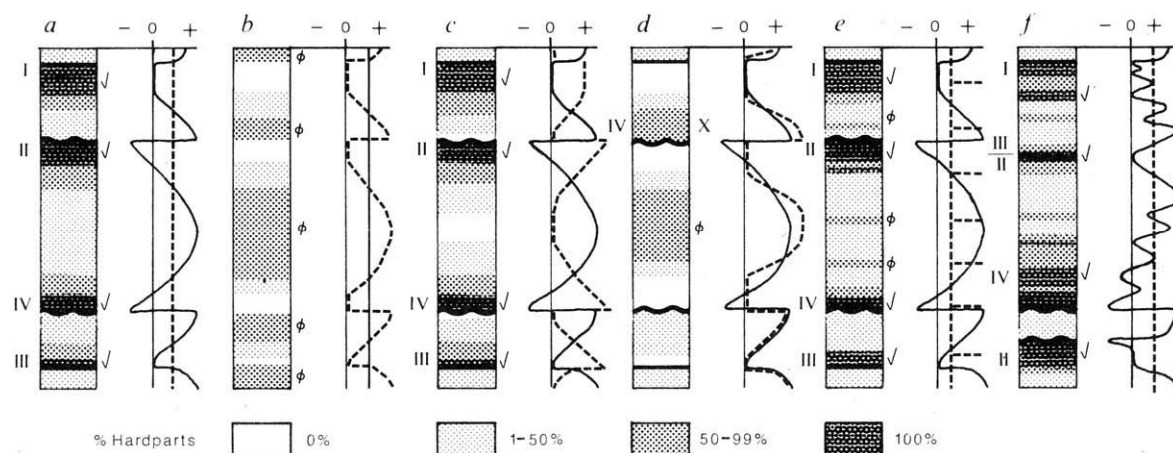


Fig. 2 Synthetic fossiliferous sequences generated from hypothetical patterns of change in sedimentation (solid curve) and hardpart input (dashed curve) provide a test of the robustness of the simple sedimentation model of fossil concentration (Fig. 1). *a*, Condition of constant hardpart input (original model conditions). The sedimentation model correctly interprets the mode of formation of all hardpart concentrations that can be classified as type I, II, III or IV (indicated by $\checkmark$). *b*, Constant sedimentation, only hardpart input varies. Concentrations form when hardpart input is relatively high, but none is associated with sharp discontinuity surfaces that permit classification by the model; the model does not lead to erroneous interpretations of the mode of formation (indicated by ϕ). *c*, Hardpart input varies inversely with sedimentation. Variation in hardpart input simply accentuates concentrations produced by excursions in sedimentation, and thus the model correctly interprets shell beds. *d*, Hardpart input varies directly with sedimentation. When the two rates are identical, fossil abundance does not vary stratigraphically (lower part of synthetic section), but relative concentrations do arise when hardpart input is allowed to exceed sedimentation. The model yields incorrect interpretations (indicated by X) of only those shell beds actually associated with discontinuities. This requires a very particular pattern of sedimentation plus perfect hydraulic equivalence of hardparts with sediment or preferential colonization by macrobenthos during the episode of maximum sedimentation. *e*, Variable net sedimentation, constant background rate of hardpart input with brief positive excursions (for example, mass mortality, allochthonous input). Synthetic pattern of fossil abundance closely mimics real stratigraphical sequences wherein some concentrations are associated with stratigraphical discontinuities, concentrations that can be classified by bed contacts are correctly interpreted by the simple sedimentation model despite random variation in hardpart input. *f*, Complex changes in sedimentation plotted against constant hardpart input. Short-term fluctuations during overall decrease or increase in net sedimentation produce internally complex fossil concentrations; back-to-back monotonic changes produce composite types I-III and II-IV deposits which share an internal discontinuity surface, and composite type III-I and IV-II beds in which both upper and lower contacts are sharp; the model is robust to any pattern of change in sedimentation.

all four large-scale shell beds (Beds 10, 14, 17 and 19 (ref. 13) in Fig. 3a) and most smaller-scale shell concentrations can be classified on the basis of bed contacts as type I, II, III or IV. These exhibit physical evidence of accumulation during intervals of reduced net sedimentation (letter-coded in Fig. 3). Expected palaeontological trends are best developed in the large-scale accumulations, which record skeletal accumulation over prolonged periods of time (estimated at 10^{-3} – 10^{-4} yr each³). As evident in the expanded columnar sections of beds 10 and 17 (Fig. 3b, c), these shell beds exhibit similar vertical sequences: a basal shell hash of largely fragmental infaunal and unbroken epifaunal hardparts; this grades into the main body of the shell bed which contains ecologically mixed and highly amalgamated assemblages of both whole and broken closely packed shells; and an upper interval of closely spaced but discrete shell horizons by which the shell bed grades into less fossiliferous overlying strata. The small-scale types I and IV shell beds which characterize the upper transitional part of these thick, complex type IV shell beds are lags of reworked hardparts from soft-bottom communities, and in some instances have been colonized by epifauna⁹. Directional shifts in bivalve morphology documented within each of the major Miocene shell beds have been interpreted as microevolutionary in origin¹⁴, but might also represent the predicted ecological response to changing substratum characteristics or to concomitant increases in water depth deepening during transgression, or the expected effects of time-averaging. Expected palaeontological trends are thus borne out on several hierarchical scales of skeletal accumulation in this particular setting⁴.

These results with both synthetic and real stratigraphical sections suggest that the model has more than heuristic value. The model should promote palaeobiological comparison of fossil assemblages because it can be used as a testable hypothesis to address a wide variety of skeletal concentrations. Hardpart accumulations can be evaluated regardless of their timescale of

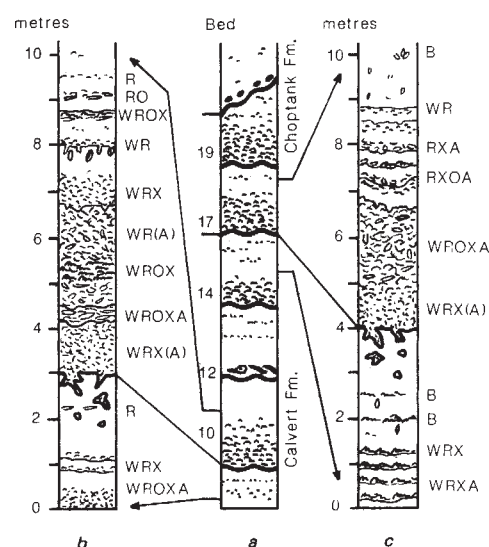


Fig. 3 The Middle Miocene Calvert and Choptank formations (composite section, middle column *a*) include four thick (1–10 m) and laterally extensive (2,500–7,000 km²) shell beds (numbered 10, 14, 17 and 19 in column *a*) and many small-scale concentrations. Most of these are associated with stratigraphical discontinuities—disconformities in the case of the four major shell beds, and scoured or winnowed bedding planes in the case of the minor accumulations (see schematic columns *b* and *c* for expanded scales)—and exhibit independent evidence of concentration during low net rates of sedimentation¹⁵. W, Winnowed matrix; R, hydraulically reoriented specimens; O, overpacked infauna; A, physically amalgamated beds; X, exhumed specimens; B, concentrations of biogenic origin, for example, clumps of shells recording concentration through biological processes, such as gregarious behaviour of the skeletal fauna or reworking by other organisms.

accumulation (for example, both rapidly formed storm beds and slowly accumulated basal lags on unconformities have type IV structures), physical dimensions, taxonomic composition, proportion of allochthonous hardparts, depositional environment and geological age. Detailed palaeoecological and evolutionary comparisons across environmental gradients or through geological time thus become feasible if a single type of shell bed is sampled throughout. For example, a transect based on erosive washover deposits, swash zone shell laminae, lags produced by channel and shoal migration, inner shelf storm concentrations and shelly proximal turbidites—which are all short-term type IV deposits—should yield assemblages that are ecologically and taphonomically more comparable with one another than to assemblages from short-term types I, II or III deposits. In addition, the absence (or otherwise divergence) of taphonomic trends from those predicted by the model permit the detection of more complex histories and conditions of fossil concentration. Successive horizons within a shell bed might be palaeontologically identical because of: very rapid concentration and burial which sequesters hardparts from destructive processes at the sea floor so that trends never develop; accumulation in environmental conditions (for example anoxia) which exclude potential colonists of dead hardparts; and homogenization by bioturbation. Very rapid formation and burial relative to rates of hardpart destruction and benthic colonization probably explain the less consistent development of expected trends within the small-scale shell beds of the Maryland Miocene, and bioturbation of shell bed contacts may well obscure the sedimentological origin of many others. The complex internal stratigraphies of the major shell beds, on the other hand, record the physical amalgamation of many small-scale concentrations into a single fossil deposit by the superposition of short-term fluctuations on a longer, overall trend in sedimentation rate, such as examined synthetically in Fig. 2f.

If sedimentation is the primary control on skeletal accumulation as indicated by the initial tests reported here, fossil concentrations can yield insights into the basic dynamics (timing and magnitude) of stratigraphical accumulation. The relative frequency of type I, II, III and IV beds should vary among depositional environments owing to differences in sediment supply and hydrodynamics, and suggests a new approach to the mapping and interpretation of sedimentary facies. Qualitatively different patterns of fossil accumulation can also be expected between transgressive and regressive phases of marine deposition, and among basins of different tectonic and latitudinal settings. By referring to fundamental rates and patterns of hardpart input and sedimentation, this approach to modelling has the potential to generate testable hypotheses for a systematic exploration of the nature of the fossil record on several scales.

I thank D. Jablonski, K. W. Flessa and P. M. Sadler for helpful comments. Research was supported by grants from the Geological Society of America, Society of Sigma Xi, Women's Seaman's Friend Society, Petroleum Research Fund of the American Chemical Society (4340-G2) and NSF EAR-8407740.

Received 17 June; accepted 28 August 1985.

1. Dodd, J. R. & Stanton, R. J. Jr *Paleoecology, Concepts and Applications* (Wiley, New York, 1981).
2. Brehrensmeier, A. K. & Kidwell, S. M. *Paleobiology* 11, 105–119 (1985).
3. Kidwell, S. M. 2nd N. Am. Paleont. Conf. Proc. 1, 295–300 (1982).
4. Kidwell, S. M. *Paleobiology* (in the press).
5. Barrell, J. *Bull. geol. Soc. Am.* 28, 745–904 (1917).
6. Campbell, C. V. *Sedimentology* 8, 7–26 (1967).
7. Tipper, J. C. *Nature* 302, 696–698 (1983).
8. Brinkmann, R. *Abh. K. Ges. Wiss. Göttingen n.F.* 13, 1–249 (1929).
9. Kidwell, S. M. & Jablonski, D. in *Biotic Interactions in Recent and Fossil Benthic Communities* (eds Tevesz, M. J. S. & McCall, P. L.) (Plenum, New York, 1983).
10. Kidwell, S. M. & Aigner, T. in *Sedimentary and Evolutionary Cycles* (eds Bayer, U. & Seilacher, A.) (Springer, Berlin, 1985).
11. Sylvester-Bradley, P. C. in *Concepts and Methods of Biostratigraphy* (eds Kauffman, E. G. & Hazel, J. E.) (Dowden, Hutchinson and Ross, Stroudsburg, 1977).
12. Williamson, P. G. *Nature* 293, 437–443 (1981); 296, 611–612 (1982).
13. Shattuck, G. B. in *The Miocene Deposits of Maryland* (eds Clark, W. B., Shattuck, G. B. & Dall, W. H.) (Maryland Geological Survey, 1904).
14. Kelley, P. H. *J. Paleont.* 57, 581–598 (1983).
15. Kidwell, S. M. *Am. Ass. petrol. Geol. Mem.* 36, 37–58 (1984).

Sex ratios of an aphid subject to local mate competition with variable maternal condition

Yuko Yamaguchi

Entomological Laboratory, College of Agriculture,
Kyoto University, Sakyo-ku, Kyoto 606, Japan

Fisher¹ was the first to argue that natural selection would adjust the sex ratio so as to equalize parental investment (PI) in the two sexes, where mating is at random. Hamilton² then showed that female-biased sex ratios would be favoured where male siblings compete for matings, a situation referred to as 'local mate competition' (LMC). In Hamilton's original model and in most subsequent LMC models, the females founding a local breeding population have equal amounts of PI to be allocated between sons and daughters. But in nature, females may differ in total PI and, as a consequence, in fecundity. Here I describe a model for the sex ratios of n co-foundresses, all of which differ in total PI. The model shows that each female with more than a specified minimum amount of PI is selected to make the same absolute investment in sons. All previously published cases of dramatic sex-ratio control occur in haplodiploid species. Here I test my theoretical model against sex-ratio data for an aphid, *Prociphilus oris*, which has a normal diploid genetic system.

Several theorists have considered the sex-ratio strategies of two co-foundresses with different clutch sizes, and they have found that the more fecund female should always produce the more female-biased ratio of investment^{3,4}. 'Double parasitism' of hosts often occurs in natural populations of parasitoid wasps, leading to local competition for mates among the offspring of two females. But in aphids, such LMC often occurs among the offspring of more than two females. Thus for aphids, a more general LMC model is needed.

The model described here embodies the following assumptions. Each local breeding population ('patch') is founded by n inseminated females selected at random from the total population at large. These females differ in their total PI: $P_1 < P_2 < \dots < P_n$, where P_i is the total PI of the i th female. Each female can control her offspring sex ratio, S_i , which is the proportion of P_i invested in sons. Mating occurs at random among the offspring born in the same patch, and all of the female offspring are successfully inseminated. Then the mated female offspring disperse, becoming thoroughly mixed with the rest of the population at large before the next generation's patches are established by taking samples of n females.

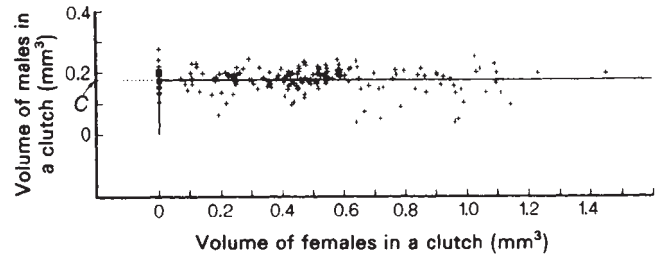
Natural selection will favour offspring sex ratios that maximize each mother's inclusive fitness. With a diploid genetic system and autosomal control of the sex ratio, mothers are related equally to their sons and daughters. At the end of the period of parental investment, the expected reproductive success of female offspring (R_f) can be taken to be a constant, equal for all daughters of all mothers. But the reproductive success of a son (R_m) depends directly and immediately on the sex ratios of the mothers in the patch. Thus the inclusive fitness of each mother is a function of her own sex ratio and of the sex ratios of the other mothers, and sex-ratio evolution takes the form of a game in which the best move for each individual depends on what all the others are doing⁵. Below, I derive the optimal sex ratios as the Nash solution or non-cooperative equilibrium of an n -person game, which in this context is the same as an evolutionarily stable strategy⁶.

The inclusive fitness of the i th mother (F_i) in a patch founded by n mothers can be written as⁷

$$F_i = r_d \frac{P_i(1 - S_i)}{w_d} R_f + r_s \frac{P_i S_i}{w_s} R_m \quad (i = 1, 2, \dots, n) \quad (1)$$

where S_i = sex ratio of the i th mother; P_i = total PI of the i th mother ($P_1 < P_2 < \dots < P_n$); r_d , r_s = relatedness of a mother to

Fig. 1 Parental investment in sons and daughters, for 176 sexuparae of the aphid *P. oriens*. Total PI in male offspring is indicated on the ordinate ($P_i S_i$), with PI in female offspring shown on the abscissa; $P_i(1 - S_i)$. Twenty sexuparae produced sons exclusively; these are represented by the vertical column of points at the left-hand end of the scattergram. Each sexupara produced at least one son. The superimposed rules show the distribution of investments predicted by the theoretical model described in the text.



daughters and sons, respectively; R_f , R_m = reproductive success of a daughter or a son, respectively; w_d , w_s = cost of a daughter or a son, respectively, in units of PI.

Assume that $r_d = r_s$ and that R_f is a constant. It follows that

$$R_m = \frac{\sum_{i=1}^n P_i(1 - S_i)/w_d}{\sum_{i=1}^n P_i S_i/w_s} R_f \quad (2)$$

From equations (1) and (2), we have

$$\frac{\partial F_i}{\partial S_i} \propto -1 + \frac{\sum_{i=1}^n P_i(1 - S_i)}{\sum_{i=1}^n P_i S_i} - \frac{\sum_{i=1}^n P_i}{(\sum_{i=1}^n P_i S_i)^2} P_i S_i \quad (3)$$

The mean sex ratio is defined as

$$\bar{S} = \frac{\sum_{i=1}^n P_i S_i^*}{\sum_{i=1}^n P_i} \quad (4)$$

where S_i^* is the optimal sex ratio of the i th mother. The Nash equilibrium conditions are

$$(\partial F_i / \partial S_i)_{S_i=S_i^*} \leq 0 \quad \text{if} \quad S_i^* = 0 \quad (5a)$$

$$(\partial F_i / \partial S_i)_{S_i=S_i^*} = 0 \quad \text{if} \quad 0 < S_i^* < 1 \quad (5b)$$

$$(\partial F_i / \partial S_i)_{S_i=S_i^*} \geq 0 \quad \text{if} \quad S_i^* = 1 \quad (5c)$$

By using equation (3), equations (5a) to (5c) can be rewritten as

$$P_i S_i^* \geq \bar{S}(1 - 2\bar{S}) \sum_{i=1}^n P_i \quad \text{if} \quad S_i^* = 0 \quad (6a)$$

$$P_i S_i^* = \bar{S}(1 - 2\bar{S}) \sum_{i=1}^n P_i \quad \text{if} \quad 0 < S_i^* < 1 \quad (6b)$$

$$P_i S_i^* \leq \bar{S}(1 - 2\bar{S}) \sum_{i=1}^n P_i \quad \text{if} \quad S_i^* = 1 \quad (6c)$$

It is convenient to define C as

$$C = \bar{S}(1 - 2\bar{S}) \sum_{i=1}^n P_i \quad (7)$$

Then equations (6a) to (6c) immediately simplify to

$$0 \geq C \quad \text{if} \quad S_i^* = 0 \quad (8a)$$

$$0 < P_i S_i^* = C \quad \text{if} \quad 0 < S_i^* < 1 \quad (8b)$$

$$0 < P_i \leq C \quad \text{if} \quad S_i^* = 1 \quad (8c)$$

I now consider where each of equations (8a) to (8c) will hold:
Case 1. $S_i^* = 0$. This is impossible. If even one mother produces any sons ($S_i > 0$), then $C > 0$ and equation (8a) is contradicted.

Case 2. $0 < S_i^* < 1$. Equation (8b) can be rearranged to give

$$S_i^* = C/P_i \quad (9)$$

and the condition $S_i^* < 1$ can be rewritten as

$$P_i < C \quad (i = 1, 2, \dots, n) \quad (10)$$

Case 3. $S_i^* = 1$. Equation (8c) implies that

$$P_i \leq C \quad (11)$$

The assumption that all mothers produce $S_i^* = 1$ leads to a contradiction, but a mixture of sex ratios from case 3 ($S_i^* = 1$) and case 2 ($0 < S_i^* < 1$) is possible, because equations (8b) and

(8c) can both be satisfied when $0 < \bar{S} < 0.5$. Thus the Nash equilibrium is

$$S_i^* = C/P_i \quad \text{for} \quad P_i > C \quad (12a)$$

$$S_i^* = 1 \quad \text{for} \quad P_i \leq C \quad (12b)$$

Next, I show that the Nash equilibrium is determined uniquely if the P_i values are given. Substituting equations (12a) and (12b) in equation (4), we obtain

$$\bar{S} = \frac{\sum_{i=1}^n \min(P_i, C)}{\sum_{i=1}^n P_i} = \left\{ \sum_{i=1}^m P_i + C(n - m) \right\} / \sum_{i=1}^n P_i \quad (13)$$

where it is assumed that the P_i values are ordered as an increasing series. Equation (13) shows that C is a semi-concave function of $\bar{S}$, while equation (7) shows that C is a convex function of $\bar{S}$ ($0 < \bar{S} < 0.5$). Thus, there is a unique solution ($\bar{S}, C$) that satisfies both equations (7) and (13), and m is uniquely determined ($P_m \leq C < P_{m+1}$).

Finally, I calculate $\bar{S}$. If all mothers produce sex ratios strictly between 0 and 1 (case 2), we have an instance of the 'special case', while if some mothers produce sex ratios of unity (case 3) we have an instance of the 'general case'. The special case represents the general case with $m = 0$.

General case.

$$S_i^* = 1 \quad (i = 1, 2, \dots, m)$$

$$S_i^* = C/P_i \quad (i = m + 1, \dots, n) \quad (14)$$

Equation (7) together with equation (13) gives

$$\bar{S} = \frac{1}{4(n - m)} \left\{ (n - m - 1) + \sqrt{(n - m - 1)^2 + 8(n - m) \sum_{i=1}^m P_i} \right\} \sum_{i=1}^n P_i \quad (15)$$

C is determined by substituting equation (15) in equation (7).
Special case.

$$S_i^* = C/P_i \quad (i = 1, 2, \dots, n) \quad (16)$$

Combining equation (6b) with equation (4) yields

$$\bar{S} = \frac{1}{2}(n - 1)/n \quad (17)$$

Substituting equation (17) in equation (7), we obtain

$$C = \left(\sum_{i=1}^n P_i / n \right) \{ (n - 1)/2n \} \quad (18)$$

The equilibrium sex ratios are uniquely determined by the P_i . The optimal sex ratio of the i th mother (S_i^*) is inversely related to her total PI (P_i), if P_i is larger than the constant C ; see equations (14) and (16). This result generalizes the existing sex-ratio models for two foundresses with unequal PI. But the most surprising result is that $C = P_i S_i^*$ —in other words, all mothers with more than C units of total PI are selected to make the same absolute investment in sons, while mothers with less than C units of PI are selected to produce sons exclusively. This is a new result.

If there is little variation in total PI, such that all mothers make some daughters (special case), then the average sex ratio

in each patch is $\bar{S} = \frac{1}{2}(n-1)/n$; see equation (17). This is the same as Hamilton's equilibrium sex ratio for the case in which all females have the same PI. But if some mothers have less than C units of PI (and therefore produce sons exclusively; general case), Hamilton's expectation does not give the exact average patch sex ratio; see equation (15).

At equilibrium, the effect of sibling competition on the inclusive fitness of each mother is the same, because each female produces the same number of sons. This would not be true if, for example, each female produced the same offspring sex ratio, regardless of her total PI.

The aphid *Prociphilus oriens* (Homoptera, Pemphigidae) is a useful species with which to test sex-ratio theories, owing to several features of its biology. It undergoes sexual reproduction once a year, in the autumn. The sexuparae ('mothers', for purposes of testing this model) produce sexuales (males and oviparous females) parthenogenetically. The sexuparae mature and feed on the secondary host, *Abies sachalinensis*. During this period, embryonic sexuales are developing within the sexuparae. When they have finished feeding, the sexuparae moult and fly to the primary host, *Fraxinus mandshurica*. They move around on the primary host for a few hours, looking for places to deposit their offspring; this wandering gives rise to a patchy distribution of sexuparae on the bark of the primary host. Within half a day the sexuparae deposit their offspring and die.

The sexuales stay at their place of birth for a few days, and then become sexually active. The males have very limited powers of movement, and they are sexually active for only about one day. Thus the sexuales produced by the mothers in a given patch are effectively a breeding population within which there is likely to be significant local mate competition. Mothers appear to have direct control over the sex ratios of their offspring⁸ and their fecundities over a wide range. The observations of mating behaviour of the sexuales showed that they cannot recognize siblings. Thus, *P. oriens* seems to satisfy all of the assumptions of the model.

Sexuales of both sexes complete their development within the bodies of the sexuparae. Lacking mouthparts, they do not feed after birth. Thus the volume of each offspring is an excellent index of its mother's investment in it⁹. The sexuales of *P. oriens* are remarkably dimorphic at birth. The average volume of a male is 0.047 mm³, while that of a female is 0.126 mm³.

Sexuparae were sampled at a site on Hokkaido, the northern island of Japan, in 1982. Sexuparae that flew to a clump of *F. mandshurica* during a period of several hours were caught as they attempted to land on the trees. They were fixed in 75% ethanol and later dissected. For each sexupara I determined the number of her male and female sexuales, and I estimated the volume of each individual sexuelle. The total PI of each sexupara (P_i) was defined as the total volume of her clutch, and her PI ratio (S_i) was then defined as the proportion of the total volume that consisted of sons.

For the sample of 176 sexuparae, the overall numerical sex ratio was almost exactly 0.5 males, while the overall PI ratio was 0.37, which is strongly and significantly female-biased. The difference between the numerical ratio and the PI ratio is caused by the sexual dimorphism of body size mentioned above.

Figure 1 shows the observed PI in sons and daughters for each sexupara. Superimposed on the data are lines illustrating the distribution of PI values predicted by the model, using the value of C calculated from the data. Mothers with less than C mm³ of total PI should invest exclusively in sons (the vertical line), while mothers with over C mm³ should invest C mm³ in sons and the rest in daughters (the horizontal line). The data cluster remarkably well about the theoretical lines.

If females could have been sampled after they had aggregated into local patches, C could have been determined directly from the model, using the observed average value of n and the observed distributions of P_i within patches. Because this was not feasible, I calculated C as the population average of the total volume of males in a clutch. This C value is equivalent to $\bar{C} = \sum_{j=1}^K C_j / K$ ($j = 1, 2, \dots, K$, where K is the number of

patches), because the total number of sexuparae sampled would divide into many patches. The good agreement of this C value with the data suggests that the variation of C between the patches is small. The observed mean investment ratio ($\bar{S} = 0.37$) implies that a typical patch contains the offspring of approximately four sexuparae; equations (15) and (17). This inferred value of n is fully consistent with my observations of the behaviour of *P. oriens* at the site from which the sexuparae were obtained. The PI distribution within a patch, however, may vary between patches in nature. Constructing a model which includes the variation of total PI between the patches and also that of the patch size, n , is a job for the future.

I thank Professor R. May for encouraging the publication of this work, and Drs Y. Iwasa, N. Yamamura and Y. Suzuki for many helpful suggestions. Professors E. Kuno and E. L. Charnov read my earlier draft and encouraged the study. I also thank Drs J. Seger and J. W. Stubblefield for their valuable comments.

Received 17 September 1984; accepted 13 September 1985.

1. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
2. Hamilton, W. D. *Science* **156**, 477-488 (1967).
3. Suzuki, Y. & Iwasa, Y. *Res. Popul. Ecol.* **22**, 366-382 (1980).
4. Werren, J. H. *Science* **208**, 1157-1159 (1980).
5. Maynard Smith, J. A. *Rev. Ecol. Syst.* **9**, 31-56 (1978).
6. Maynard Smith, J. & Price, G. R. *Nature* **246**, 15-18 (1973).
7. Oster, G. F., Eshel, I. & Cohen, D. *Theor. Popul. Biol.* **12**, 49-85 (1977).
8. Yamaguchi, Y. in *Proc. 18th int. Ethological Conf.* 313 (Brisbane, 1983).
9. Trivers, R. L. in *Sexual Selection and the Descent of Man 1871-1971* (ed. Campbell, R.) (Aldine, Chicago, 1972).

Age-impaired impulse flow from nucleus basalis to cortex

Gary Aston-Jones, Joseph Rogers*, Robert D. Shaver, Timothy G. Dinan† & Donald E. Moss‡

Department of Biology, Rm 1009 Main Bldg, Washington Square Center for Neural Science, New York University at New York, New York 10003, USA

* Department of Neurology, University of Massachusetts Medical School, 55 Lake Avenue, Worcester, Massachusetts 01605, USA

‡ Department of Psychology, University of Texas, El Paso, Texas 79968, USA

Recent studies have renewed interest in the role of acetylcholine (ACh) in the cognitive changes associated with ageing and dementia¹⁻⁴. Deficits in cortical choline acetyltransferase (ChAT) in Alzheimer's disease have been consistently demonstrated⁵⁻⁹, while other research has suggested a connection between deterioration of cortical ACh fibres and dementia⁴. However, despite clear biochemical and anatomical evidence for a fall in ACh in dementia¹⁻⁹, results of therapeutic trials with cholinergic agonists, precursors and cholinesterase inhibitors have been inconsistent⁶⁻⁹. Such findings suggest that cortical cholinergic disorders are not wholly a function of simple biochemical change; alterations of impulse flow along cholinergic fibres could well be as debilitating. An important extrinsic source of cortical ACh innervation derives from neurones diffusely located in rat basal forebrain¹⁰⁻¹⁶, denoted the nucleus basalis (NB)^{15,17}. We have now investigated the impulse conduction properties of cortically projecting, putatively cholinergic NB axons in adult and aged rats and have found that conduction latencies from NB to frontal cortex are significantly longer (by 51%) in aged animals. In addition, systematic analysis varying cortical stimulation depth revealed that these longer latencies are due entirely to decreased conduction velocities in the subcortical fibre projections. Indeed, intracortical velocities were virtually identical in the two groups. Our results indicate that ageing occasions a decrease in the temporal fidelity of impulse flow in the cholinergic input to the cortex from the NB, a previously overlooked but potentially important element in cognitive deficits that occur with age.

† Present address: Department of Psychological Medicine, St Bartholomew's Hospital, West Smithfield, London EC1A 7BE, UK.

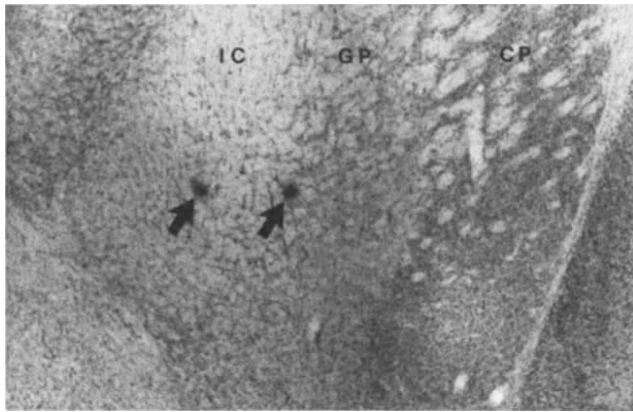


Fig. 1 Photomicrograph of a coronal section through the basal forebrain of an aged experimental rat (neutral red stain). Pontamine sky blue (spots in NB area, at arrows) was iontophoresed at the locations of two cells driven antidromically from prefrontal cortex. All cells in this study were localized histologically to the NB area in this manner. For calibration, Pontamine spots are 500 μm apart. GP, globus pallidus; IC, internal capsule; CP, caudate putamen.

Six adult (6–8 months) and seven aged (20–29 months) female Sprague-Dawley rats were anaesthetized with chloral hydrate throughout experimental sessions. Aged rats were in good health and remained as viable as adult rats under anaesthesia. A twisted pair of 250- μm -diameter wires, insulated except for bluntly cut tips, was used to stimulate the prefrontal cortex, while a glass micropipette served to record impulses from single NB neurones. The stimulation electrode was placed 0.5 mm below the cortical surface and then raised or lowered in 0.3-mm steps when testing individual cells. Details of stimulation and recording procedures are published elsewhere^{18,19}. Recording sites were marked at the end of successful penetrations by iontophoresing dye from pipette tips. All cells reported here were histologically localized to the NB area (Fig. 1), corresponding to previous anatomical maps of cortically projecting acetylcholinesterase or ChAT-positive neurones^{13,15,20,21}. To control for possible effects of decreased ovarian function with age, half the control adults were ovariectomized 4–6 weeks before experiments. No difference was seen between ovariectomized and intact control rats, and these results were pooled.

Of more than 400 neurones recorded in the vicinity of the NB, 35 in aged animals and 25 in young animals were activated antidromically from the prefrontal cortex. Criteria for antidromicity included constant latency driving, high-frequency activation and collision between spontaneous and driven impulses^{18,19} (Fig. 2). These cells exhibited a variety of rates and patterns of spontaneous activity, resembling the physiological heterogeneity previously found for cortically projecting NB neurones^{18,19}. As also reported previously for young adult male rats^{18,19}, conduction latencies varied widely, yielding values ranging from 1.2 to 16.0 ms in young controls. Despite this heterogeneity, latencies in aged rats were longer overall than in controls, ranging from 2.1 to 27 ms. The mean latency for aged subjects (8.6 ± 0.5 ms) was significantly greater than that for the young group (5.7 ± 0.3 ms), exhibiting a 51% increase ($F = 21.1$, $P < 0.001$).

Nearly every NB neurone, in both young and old rats, could be activated antidromically from the entire depth of the cortex, thus enabling us to compare intracortical and subcortical aspects of the age change in NB axon conduction velocity. Antidromic latencies decreased significantly in a linear fashion overall as the stimulating electrode was lowered to deeper cortical sites ($F = 5.80$, $P < 0.02$) (Fig. 3a). However, although conduction latencies for old rats were consistently longer than for young rats at all cortical depths, slopes of the curves were virtually identical for both groups, yielding estimated intracortical conduction velocities of 0.8 m s^{-1} regardless of age (Fig. 3a). Thus, the entire difference in conduction times between the two groups

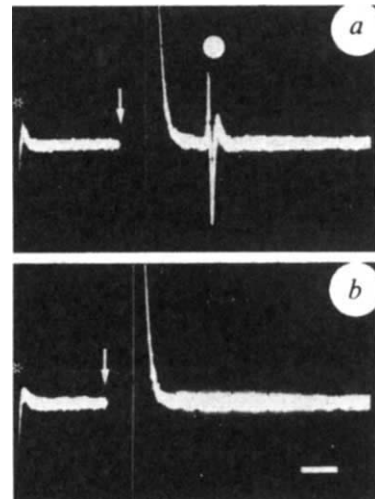


Fig. 2 Oscilloscope photographs illustrating collision test for an NB neurone driven from prefrontal cortex in an aged rat. *a*, Stimulation of cortex (arrow) 12 ms after spontaneous spikes (asterisk) elicits driven spikes (filled circle) at 10 ms latency. *b*, Driven spikes are occluded for similar stimuli (arrow) delivered 11 ms after spontaneous spikes (asterisk), indicating collision between spontaneous and driven impulses. Horizontal calibration, 5 ms. Ten superimposed sweeps in each trace.

is accounted for by significantly slower subcortical conduction velocities in aged rats (Fig. 3b).

In contrast to the above conduction changes, several other physiological parameters of NB neurones remained stable in old age. (1) Absolute refractory periods averaged 2.6 ± 0.3 ms in adult controls and 2.4 ± 0.2 ms in aged rats. (2) The number of cells that could be activated antidromically averaged 0.8 ± 0.2 cells per penetration for both age groups. (3) The propensity of axons to exhibit multiple discrete antidromic latencies, previously suggested to reflect intracortical branching¹⁹, was equally evident in both age groups, with averages of 2.8 ± 0.4 and 2.4 ± 0.3 discrete latencies per cell in young and old animals, respectively. (4) Thresholds for antidromic activation were not significantly different between age groups across cortical stimulation depths ($P > 0.1$, Student's *t*-test). (5) There was no apparent difference in the amplitudes of driven impulses in young and aged rats.

Several mechanisms may underlie these age-dependent decreases in NB conduction velocity. (1) Selective loss of larger NB axons, perhaps by shrinkage²², could produce such an effect. However, this process would seem likely also to alter stimulation thresholds, refractory periods and intracortical conduction velocities, none of which showed any significant change with age. Also, the distribution of antidromic latencies revealed an overall shift to longer latencies with age, rather than a loss of one subpopulation. These results, together with the finding that the frequency of encountering driven cells does not change with age, indicate that fibre shrinkage or selective cell loss in aged subjects is unlikely to account for our findings. (2) Similar reasoning argues against the possible mediation of these changes in conduction velocity by altered biophysical properties of the NB axolemma. Although age-associated changes such as increased membrane resistance and viscosity have been reported for other systems^{23,24}, such mechanisms would again be expected to alter not only conduction velocity, but also other electrophysiological characteristics. (3) A third possibility, more consistent with the present data, is demyelination of NB axons. The subcortical conduction velocities obtained for young animals (most of which were $>3.0 \text{ m s}^{-1}$) indicate that NB axons are normally myelinated subcortically. However, conduction velocities for NB axons within cortex are significantly slower, with mean values (0.8 m s^{-1}) more suggestive of non-myelinated axons^{25,26}. Similar results have been obtained for adult male rats¹⁹. Thus, like mossy fibre afferents to the cerebellar cortex²⁷, NB fibres

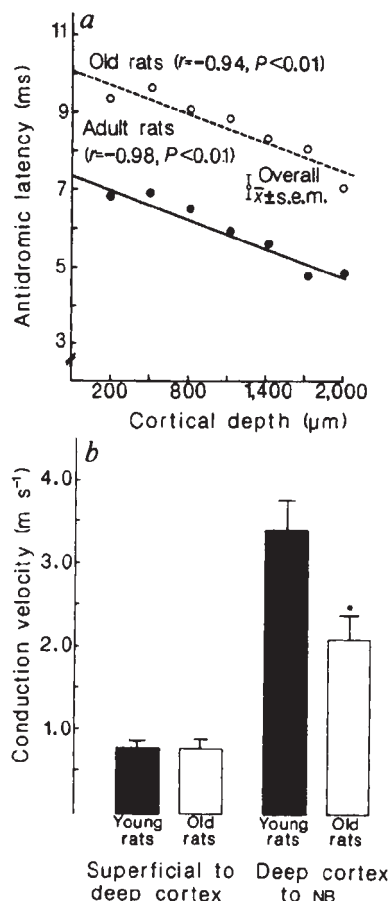


Fig. 3 *a*, Mean antidromic conduction latency for NB neurones in adult (filled circles) and aged (open circles) subjects after threshold stimulation at various depths in the prefrontal cortex, as indicated on the abscissa. Note that although conduction latencies are longer at all depths in aged rats, the slopes of the curves are virtually identical, indicating similar conduction times in the two age groups for intracortical impulse flow. In repeated measures analysis of variance, the interaction of age group with cortical depth was nearly zero, indicating that the effect of age was consistent at all cortical depths. The high correlation coefficient (r values) obtained in the regression analysis indicates that these fibres have a strong radial component in their intracortical trajectories, supporting previous anatomical studies^{11,12,36-38}. *b*, Mean conduction velocities for NB cortical afferent fibres in adult (filled bars) and aged (open bars) rats. Conduction velocities within the cortex were calculated for each cell and then averaged, using the difference between the latencies at threshold intensities for 200- and 2,000-μm stimulation depths and a corresponding distance of 1.8 mm. Subcortical conduction speeds (deep cortex to NB) were similarly calculated for each cell using the threshold latency at 2,000 μm and an estimated distance from NB to cortex of 10 mm. Note that conduction velocities within the cortex are virtually identical for adult and aged subjects. Subcortical conduction velocities for young rats (mean $\pm$ s.e.m. = 3.4 ± 0.4 m s⁻¹) are over fourfold higher than those in the cortex, and show a significant decline with ageing (conduction velocity for aged rats = 2.1 ± 0.3 m s⁻¹; $t_{38} = 2.43$, $P = 0.02$).

may be myelinated until they enter the cortical grey matter¹⁹. Given such a subcortically myelinated projection system, the hypothesis of age-related demyelination is well in accord with all aspects of our data. It explains, for example, why intracortical conduction velocities remain stable with age while subcortical velocities decline markedly. Similarly, demyelination subcortically would not require alteration in refractory periods or thresholds within the cortex. Finally, such a mechanism would leave the relative numbers of fibres and cells intact, a point consistent with our results.

Further research, particularly at ultrastructural levels, will be necessary to test our hypothesis. We note, however, that de-

myelination, whether as a neuropathy in itself or as an effect secondary to axonal pathology, is an established basis for neurological disorders, occurring, for example, in Charcot-Marie-Tooth disease and Friedreich's ataxia²⁸.

Ours is not the first report of altered fibre conduction velocity in senescence. Decreased impulse speeds have been found in human peripheral²⁹⁻³¹ and spinal nerves³¹, as well as rodent motoneurons²³ and cerebellar parallel fibres³². However, examples where there is no senescent change in fibre conduction velocity have been reported in both central^{33,34} and peripheral³⁵ nervous systems. Thus, decreased impulse conduction velocity in old age occurs in a restricted set of fibre systems. This selective action may underlie distinct symptoms of the overall ageing syndrome. For example, diminished cerebellar parallel fibre conduction velocity may impair motor coordination³², a common problem in ageing. Similarly, reduced temporal fidelity in the NB cortical afferent system, as described here, may promote cognitive deficiencies characteristic of senescence, and especially pronounced in dementia of the Alzheimer's type.

Although many behavioural studies of ageing have proposed a relationship between cognitive deficits and changes in cholinergic systems⁶⁻⁹, evidence for the latter at the biochemical level is scant. Cortical ChAT decline, for example, occurs consistently in Alzheimer's disease, but not in normal ageing¹⁹. This seldom-addressed paradox may be explained by our findings. Demyelination of cortically projecting cholinergic fibres with age would not necessarily require cholinergic biochemical concomitants, but would nonetheless impair the temporal fidelity of NB-cortical processing, thereby disrupting normal cognitive mechanisms.

If the primary function of the nervous system is to process information by virtue of electrical impulse flow, physiological measurements such as those in the present study may provide a relevant and sensitive index of geriatric cognitive decline.

This work was supported by PHS grants NS 19360 and NS 22320 (G. A.-J.), BRSG grants SO7RR7149-09 (G. A.-J.) and 6-32840 (J.R.), and by the Alzheimer's Disease and Related Disorders Association (G. A.-J. and J.R.).

Received 15 April; accepted 3 October 1985.

- Price, D. L. *et al. Neurosci. Comments* **1**, 84-92 (1982).
- Whitehouse, P. J. *et al. Science* **215**, 1237-1239 (1982).
- Coyle, J. T., Price, D. L. & DeLong, M. R. *Science* **219**, 1184-1190 (1983).
- Struble, R. G., Cork, L. C., Whitehouse, P. J. & Price, D. L. *Science* **216**, 413-415 (1982).
- McGeer, E. G. & McGeer, P. L. in *Advances in Behavioral Biology: Neurobiology of Aging* (eds Ord, J. M. & Brizze, K. R.) 287-305 (Plenum, New York, 1975).
- Corkin, S. *Trends Neurosci.* **4**, 287-290 (1981).
- Kubanis, P. & Zornetzer, S. F. *Behav. neural Biol.* **31**, 115-172 (1981).
- Bartus, R. T., Dean, R. L., Beer, B. & Lippa, A. S. *Science* **217**, 408-417 (1982).
- Rogers, J. & Bloom, F. E. in *Handbook of the Biology of Aging* (eds Schneider, E. & Finch, C. E.) 645-691 (Van Nostrand Reinhold, New York, 1985).
- Schute, C. C. D. & Lewis, P. R. *Brain* **90**, 497-520 (1967).
- Johnston, M. V., McKinney, M. & Coyle, J. T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5392-5396 (1979).
- Johnston, M. V., McKinney, M. & Coyle, J. T. *Expl Brain Res.* **43**, 159-172 (1981).
- Lehmann, J., Nagy, J. I., Atmadja, S. & Fibiger, H. C. *Neuroscience* **5**, 1161-1174 (1980).
- Wenk, H., Bigl, V. & Meyer, U. *Brain Res. Rev.* **2**, 295-316 (1980).
- Bigl, V., Woolf, N. J. & Butcher, L. L. *Brain Res. Bull.* **8**, 727-749 (1982).
- McKinney, M., Coyle, J. T. & Hedreen, J. J. *Comp. Neurol.* **217**, 103-121 (1983).
- Parent, A., Gravel, S. & Olivier, A. *Adv. Neurol.* **24**, 1-11 (1979).
- Aston-Jones, G., Shaver, R. & Dinan, T. *Neurosci. Lett.* **46**, 19-24 (1984).
- Aston-Jones, G., Shaver, R. & Dinan, T. *Brain Res.* **325**, 271-285 (1985).
- Rye, D. B., Wainer, B. J., Mesulam, M.-M., Mufson, E. J. & Saper, C. B. *Neuroscience* **13**, 627-643 (1984).
- Woolf, N. J., Eckenstein, R. & Butcher, L. L. *Brain Res. Bull.* **13**, 751-784 (1984).
- Pearson, R. C. A. *et al. Brain Res.* **289**, 375-379 (1983).
- Chase, M. H., Morales, F. R., Boxer, P. A. & Fung, S. J. *Expl Neurol.* (in the press).
- Heron, D. S., Hershkowitz, M., Shinitzky, M. & Samuel, D. in *Neurotransmitters and their Receptors* (eds Littauer, U., Didal, Y., Silman, I., Peichberg, V. & Vogel, Z.) 125-137 (Wiley, New York, 1980).
- Waxman, S. G. & Bennette, M. V. L. *Nature new Biol.* **238**, 217-219 (1972).
- Swadlow, H. A. & Waxman, S. G. *Expl Neurol.* **53**, 128-150 (1976).
- Barr, M. L. *The Human Nervous System* 2nd edn, 156-158 (Harper & Row, Hagerstown, 1974).
- Griffin, J. W. in *The Clinical Neurosciences* (ed. Rosenberg, R. N.) 538-539 (Churchill-Livingstone, New York, 1983).
- Wagman, I. H. & Lesse, H. J. *Physiol. Lond.* **15**, 235-244 (1952).
- Behse, F. & Buchthal, F. J. *Neurol. Neurosurg. Psychiatr.* **34**, 404-414 (1971).
- Dorfman, L. J. & Bosley, T. M. *Neurology* **29**, 38-44 (1979).
- Rogers, J., Zornetzer, S. F. & Bloom, F. E. *Neurobiol. Ageing* **2**, 15-25 (1981).
- Landfield, P. W. & Lynch, G. J. *Gerontol.* **32**, 523-533 (1977).
- Landfield, P. W., McGaugh, J. L. & Lynch, G. J. *Brain Res.* **150**, 85-101 (1978).
- Smith, D. O. & Rosenheimer, J. L. *Expl Neurol.* **83**, 358-366 (1984).
- Krnjevic, K. & Silver, A. J. *Anat.* **99**, 711-759 (1965).
- Hedreen, J. C. *et al. J. comp. Neurol.* **226**, 246-254 (1984).
- Houser, C. R., Crawford, G. D., Barber, R. P., Salvaterra, P. M. & Vaughn, J. E. *Brain Res.* **266**, 97-119 (1983).

A late-differentiation antigen associated with the helper inducer function of human T cells

Nicole Suciu-Foca*, Elaine Reed*, Pablo Rubinstein*, Wilma MacKenzie*, Ah-kau Ng* & Donald W. King†

*Columbia University, College of Physicians and Surgeons, 630 West 168th Street, New York, New York 10032, USA

†The University of Chicago, Pritzker School of Medicine, 950 East 59th Street, Chicago, Illinois 60637, USA

T lymphocytes possessing helper function produce soluble factors that greatly augment B-cell proliferation and differentiation into antibody-secreting cells¹. In humans the subset of T lymphocytes bearing the T4 surface antigen comprises most of the cells that display helper activity and recognize class II antigens of the major histocompatibility complex (MHC), while the subset bearing the T8 antigen comprises T cells recognizing class I MHC antigens and exhibiting cytotoxic or suppressor function¹⁻³. Monoclonal antibodies to T4 or T8 greatly inhibit the cognitive and effector function of cells with the corresponding phenotype¹⁻⁴. This function/phenotype correlation is not absolute, however, for there are many examples of T8-positive clones that recognize MHC class II antigens and have helper activity, as well as of T4-positive clones with suppressor or cytotoxic function⁵⁻⁹. Recently a family of cell-surface neoantigens, which might be relevant to T-cell function and which are present on activated but not on resting T lymphocytes, has been identified in mouse and humans using monoclonal antibodies¹⁰. Some of these antibodies block the cytolytic activity of alloreactive T-cell clones, suggesting the possible involvement of such molecules in the activation of cytotoxic T-cell clones or in the lytic process itself¹⁰⁻¹³. We now describe a similar late-differentiation antigen (LDA₁) that is expressed by human T lymphocytes only following activation and is recognized by a monoclonal antibody that inhibits the antibody-inducing helper function of T lymphocytes.

Monoclonal antibody anti-LDA₁ was produced by immunizing a BALB/c mouse with an alloreactive T-cell clone (TCC 19) possessing helper function^{14,15}. This hybridoma was originally selected because it reacted with the immunizing T-cell clone (no. 19) but not with peripheral blood T or B lymphocytes or with Epstein-Barr virus-transformed B-lymphoblastoid cell lines from the same donor or from 30 unrelated individuals. Further screening of this antibody on alloresponsive T-cell clones showed that it reacted with all of the 24 helper but with none of the 14 cytotoxic T-cell clones tested. This antibody is of IgG1 subclass as determined by standard immunodiffusion procedures, and was used as an optimally diluted ascitic fluid in most experiments.

Indirect immunofluorescence using an Ortho Spectrum III cytofluorograph^{14,15} showed that LDA₁ is absent on resting T lymphocytes, but appears 1-3 days after the peak of DNA synthesis on a fraction of T lymphoblasts stimulated *in vitro* with plant mitogens, soluble antigens or allogeneic cells. The expression of LDA₁ is dependent on cell proliferation and protein synthesis, as it is prevented by treatment of cells with mitomycin C or emetine-HCl shortly after activation¹⁶. Anti-LDA₁ had no effect on the blastogenic responses of fresh T lymphocytes to plant mitogens, soluble antigens [tetanus toxoid (TT) and purified protein derivative of tuberculin (PPD)] or allogeneic stimulating cells.

Cytofluorometric evaluation of LDA₁ expression on cells grown in primary mixed lymphocyte culture (MLC) showed that 30% LDA₁-positive T blasts are present by day 9, decreasing to <5% by day 14. The fraction of the population expressing LDA₁ was increased, however, to 70% when day-6 MLC-T lymphoblasts were expanded by continuous stimulation with irradiated cells from the sensitizing donor in medium supplemented with interleukin-2.

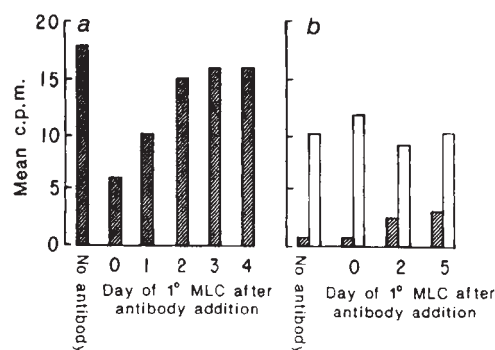


Fig. 1 Replicate cultures containing 10^7 PBM from an individual with the HLA-DR5,9 phenotype were stimulated in 20-ml tissue culture flasks with an equal number of irradiated PBM from an HLA-DR5,7 donor. Anti-LDA₁ was added at a final dilution of 1:100 to individual culture flasks on day 0, 1, 2, 3, 4 or 5. Control cultures contained no antibody. Ten days after the initiation of the primary cultures, cells were collected from each flask, centrifuged on a Ficoll gradient for removal of dead cell debris, adjusted to 5×10^4 viable responding cells per Microtest well and challenged in triplicate reactions with HLA-DR7-positive (■) or DR7-negative (□) stimulating cells. Secondary cultures were labelled with ^3H -thymidine for 18 h and collected on day 3 (a) or day 5 (b) following restimulation. Results are expressed as mean c.p.m. $\times 10^3$.

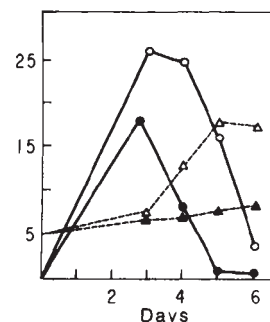


Fig. 2 The NSF-K T-cell line was developed by priming PBM from a DR2 homozygous individual to irradiated PBM from a DR3 homozygous stimulator. Day-6 MLC blasts were propagated for 28 days by weekly stimulation with DR3-positive cells in medium containing 20% interleukin-2. Responding T lymphoblasts (5×10^6 culture) were tested for reactivity to DR3-positive stimulating cells in the absence or presence of anti-LDA₁ (1:100 final dilution). The number of viable cells and the rate of ^3H -thymidine incorporation were determined on days 3-6. ○, c.p.m. $\times 10^3$ per 10^5 T cells grown without antibody; ●, c.p.m. $\times 10^3$ per 10^5 T cells grown with anti-LDA₁; Δ, number of cells ($\times 10^6$) in medium without antibody; ▲, number of cells ($\times 10^6$) in medium containing anti-LDA₁.

To determine whether the number of LDA₁-positive cells could be increased further, T-cell lines that had been propagated for 1 month *in vitro* and which expressed LDA₁ on 70% of the cells, were sorted into LDA₁-positive and -negative fractions using a fluorescence-activated cell sorter (FACS-I; Becton-Dickinson). Both the positively and negatively selected fractions exhibited $30 \pm 12\%$ LDA₁-positive T lymphoblasts after 48 h, and $60 \pm 15\%$ positive cells 120 h after sorting and expansion in cultures. Similar variations in LDA₁ expression were found on the immunizing TCC 19. This helper line, whose monoclonal origin was confirmed by studying the rearrangement of the T-cell antigen receptor β -chain¹⁷, displayed $40 \pm 18\%$ and $80 \pm 15\%$ LDA₁-positive cells, respectively, at 48 and 120 h following each feeding. This indicates that LDA₁ expression is modified on the surface of the alloactivated T lymphoblasts by which it is synthesized.

When added at the initiation of primary MLC cultures, the anti-LDA₁ antibodies did not significantly inhibit the primary day-5 blastogenic response, nor the generation of cytotoxic T

cells in primary MLC or the lytic activity of such cells as determined in a ^{51}Cr -release assay¹⁵.

Because the expression of LDA₁ is acquired only during late stages of the primary MLC response (day 9), we investigated the effect of anti-LDA₁ on secondary MLC reactions, which are usually tested on day 10. The secondary MLC response or the primed lymphocyte test detects accelerated memory responses displayed by cells that have been primed to a specific class II MHC antigen in primary cultures¹⁸. Non-activated cells from the same cultures do not participate in the secondary MLC response, yet maintain the ability to display primary blastogenic responses to the HLA-D/DR antigens for which they bear specific receptors.

Lymphocytes from a responder bearing the HLA-DR5, 9 phenotype were primed in MLC to HLA-DR5, 7 stimulating cells. Anti-LDA₁ was added to individual cultures on day 0, 1, 2, 3, 4 or 5 and the secondary MLC response was tested on day 10 by rechallenging the cultures with stimulating cells. The reaction to specific, HLA-DR7-positive stimulating cells was an accelerated memory response, which peaked at day 3 and declined by day 5. The day-3 response was strongly inhibited when anti-LDA₁ was added to the cultures within the first 24 h of primary stimulation, but not when added later (Fig. 1a). None of the cultures showed accelerated (day 3) memory responses to HLA-DR7-negative stimulating cells. However, they displayed primary-type blastogenic responses to such cells on day 5, as expected for a polyclonal population of T cells that can recognize any allelic variants of (non-self) HLA-D/DR antigens in primary cultures. This primary anti-HLA-DR blastogenic response was not significantly inhibited in cultures grown with anti-LDA₁ (Fig. 1b). It appears, therefore, that anti-LDA₁ inhibits the capacity of primed T cells to respond to the sensitizing antigen, as further suggested by its suppressive effect on alloreactive T-cell lines (Fig. 2). The finding that LDA₁ is expressed by T lymphocytes only after activation may explain the difference between the effect of anti-LDA₁ on primary and secondary MLC.

Table 1 Effect of anti-LDA₁ on T-helper function

Individual	Antibody added	Immunoglobulin (ng ml ⁻¹) in cultures with					
		No antigen		Tetanus toxoid		PWM	
		IgG	IgM	IgG	IgM	IgG	IgM
1	None	120	70	360	460	2,400	2,020
	LDA ₁	130	60	130	100	410	140
	MB 40.2	110	120	420	510	2,570	2,100
	L243	120	90	120	100	560	180
2	None	160	40	700	2,240	1,050	1,940
	LDA ₁	200	20	50	100	300	110
	MB 40.2	130	140	650	2,300	1,160	2,010
	L243	170	30	110	140	250	100
3	None	180	160	1,100	2,800	2,600	3,460
	LDA ₁	150	110	380	630	850	720
	MB 40.2	110	120	1,050	2,700	2,580	3,500
	L243	140	150	250	470	760	680

PBM (1×10^5 per reaction) from three individuals recently sensitized to TT were cultured in a round-bottomed microculture plate at 37 °C in a humidified 5% CO₂ atmosphere in RPMI 1640 medium supplemented with 15% heat-inactivated fetal calf serum (FCS), 200 mM L-glutamine, 25 mM HEPES buffer and 1% penicillin/streptomycin. PWM (Gibco) was added to the cultures to a final concentration of $10 \mu\text{g ml}^{-1}$. TT (Massachusetts Public Health Biological Laboratories) was used at a final dilution of 1:400. Cultures were tested in parallel without antibody, with anti-LDA₁, with MB 40.2, which is an IgG1 antibody reacting with the responders HLA-B antigens, or with L243, which is of IgG2A subclass and reacts with human Ia. The amount of human IgG and IgM contained in 50- μl aliquots of day-6 culture supernatant was quantitated in round-bottomed polyvinyl trays (Costar) coated with rabbit anti-human IgG or IgM (Dako Accurate Chemicals). Known amounts of human IgG or IgM in the same medium were tested as standards. Triplicate wells were used for each reaction. Following 2 h of incubation, wells were emptied, washed in 1% FCS in phosphate-buffered saline (PBS) and covered with 50 μl of a 1:3,000 or 1:1,000 dilution of peroxidase-conjugated rabbit anti-human IgG or IgM respectively. Trays were incubated for 1 h at room temperature in the dark, then washed five times in 1% FCS/PBS. 2,2-Azino-di-(3-ethylbenzthiazoline) sulphonic acid (50 μl), diluted 1:100 in 0.1 M citrate buffer (pH 4.2) containing 0.03% hydrogen peroxidase, was added to each well. After 30 min of incubation at room temperature in the dark, 50 μl of a 1% solution of SDS was added to stop the reaction. The absorbance at 405 nm was read with an automated photometer (Flow Multiskan).

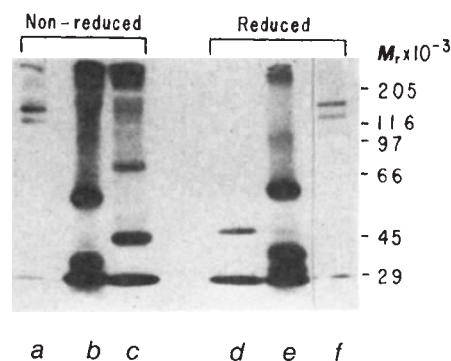


Fig. 3 SDS-polyacrylamide gel electrophoresis (PAGE) of ^{125}I -labelled membrane proteins of TCC 19 cells. 2×10^7 cells were surface labelled with 1 mCi of Na^{125}I (Amersham) using the lactoperoxidase technique as previously described¹⁴. The labelled cells were washed and lysed in 0.5 ml of 100 mM Tris-HCl, pH 7.4/1 mM phenylmethylsulphonyl fluoride/1 mM EDTA/0.5% Triton X-100. The cell lysate was centrifuged at 40,000g for 30 min and the supernatant was filtered through a 0.2- μm Millipore filter. The filtered supernatant was precleared three times by sequential 1-h incubation at 4 °C with protein A-Sepharose, rabbit anti-mouse IgG-coupled Sepharose and an irrelevant monoclonal antibody coupled to Sepharose. The precleared cell lysate supernatant was incubated for 4 h with specific antibody coupled to Sepharose-4B, following the immunoprecipitation procedure described before¹⁴. Immune precipitates were then processed for SDS-PAGE in the presence (reducing condition) of 5% 2-mercaptoethanol using a 7.5% polyacrylamide gel. Lanes a, f, anti-LDA₁; lanes b, e, antibody A3 to a framework determinant of human MHC class II antigen; c, d, antibody NAM 1 to human β_2 -microglobulin.

T cells recognizing class II MHC antigens comprise the population of lymphocytes which produce helper factors required for the differentiation of activated B cells into antibody-producing cells. To determine whether T-helper function is also affected by anti-LDA₁, we tested its effect on immunoglobulin production in cultures in which peripheral blood mononuclear cells (PMB) had been primed to TT or pokeweed mitogen (PWM). Following PWM stimulation, $30 \pm 5\%$ of PWM T blasts expressed LDA₁ by day 4 and $60 \pm 8\%$ were positive by day 6. The equivalent numbers for TT blasts were 19 ± 6 and 42 ± 7 , respectively. Anti-LDA₁ had a strong inhibitory effect on immunoglobulin production in PBM cultures stimulated with PWM or TT (Table 1). Amounts of human IgG and IgM present in the day-6 supernatant of such cultures were significantly lower than those found in cultures grown without monoclonal antibodies or with monoclonal antibody MB 40.2 (ref. 19), which reacts with HLA-B antigens and was used as an isotype-matched control.

The degree of inhibition produced by anti-LDA₁ was of the same order of magnitude as that produced by monoclonal antibody L243 (ref. 20), which reacts with Ia antigens on B cells, monocytes and activated human T cells. Similarly, anti-LDA₁ was as efficient as L243 in inhibiting the capacity of helper T-cell clones possessing anti-HLA-DR¹⁵ or anti-PPD reactivity²¹ to induce B lymphocytes to produce antibodies (Table 2), but anti-LDA₁ had no effect on antibody production by activated, Epstein-Barr virus-transformed β -lymphoblastoid cell lines. Thus, for example, levels of human IgM produced by 1×10^3 B lymphoblasts from one such line in 24, 48, 72 and 96 h were 62, 114, 226 and 350 ng ml⁻¹ in cultures without antibody and 60, 140, 220 and 380 ng ml⁻¹ in cultures containing anti-LDA₁. Immunoprecipitation studies of the molecule recognized by anti-LDA₁ were performed using ^{125}I -labelled extracts from the immunizing TCC 19. Two bands of ~116,000 and 150,000 relative molecular mass (M_r) were found in reducing and nonreducing conditions (Fig. 3).

Our data indicate that anti-LDA₁ detects a late-differentiation antigen expressed by activated T lymphoblasts and inhibits their

Table 2 Effect of anti-LDA₁ on antibody induction by helper T-cell clones

	Immunoglobulin (ng ml ⁻¹) in cultures containing:					
	No antibody		Anti-LDA ₁		L243	
	IgG	IgM	IgG	IgM	IgG	IgM
Anti-HLA-DR1	164	135	30	19	25	18
Anti-PPD	184	98	42	26	38	15

Alloreactive T-cell clones were tested for helper function by incubating 1×10^4 cloned cells with 5×10^4 stimulator type B cells in RPMI 1640 medium containing FCS, glutamine antibodies and PWM, as described previously¹⁵. PPD-specific T-cell clones (1×10^4 cells per reaction) were incubated with autologous T-lymphocyte-depleted PBM in medium containing PPD. Triplicate cultures with no antibody, with anti-LDA₁ or with L243 were tested in parallel. Supernatants were collected from day-6 cultures and the amount of human IgG and IgM was quantitated by enzyme-linked immunosorbent assay. The amount of IgG and IgM produced by B cells alone was less than 25 ng ml^{-1} and was subtracted from all values.

capacity to produce or deliver the helper signal(s) required for the differentiation of antibody-producing B cells. The inhibitory effect of this antibody is as strong as that displayed by anti-Ia antibodies which interfere with T-B interaction. LDA₁ may represent a cell-surface protein encoded by a 'helper' function gene or by a T-cell receptor gene (possibly Fc) involved in regulating antibody production.

Isolation of the gene and determination of the relevance of its product in autoimmune conditions, allograft rejection or T-cell malignancies is of obvious interest for the future.

Received 6 August; accepted 14 October 1985.

- Hood, L. E., Weissman, J. L., Wood, W. B. & Wilson, J. H. (eds) in *Immunology*, 259-314 (Benjamin/Cummings, Menlo Park, 1984).
- Reinherz, E. L. & Schlossman, S. F. *Cell* **19**, 821-827 (1980).
- Reinherz, E. L., Kung, P., Goldstein, G. & Schlossman, F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4061-4065 (1979).
- Littman, D. R., Thomas, Y., Maddon, P. J., Chess, L. & Axel, R. *Cell* **40**, 237-246 (1985).
- Jeannot, M. & Chardonnet, X. *Transplant. Proc.* **17**, 714-715 (1985).
- Flomenberg, N., Duffy, E. & DuPont, B. *Scand. J. Immun.* **19**, 237-245 (1984).
- Rohowsky, C., Kung, P., King, D. W. & Suci-Foca, N. *Hum. Immun.* **9**, 103-110 (1984).
- Suci-Foca, N. *et al. Fedn Proc.* **44**, 107-109 (1985).
- Davignon, D., Martz, E., Reynolds, T., Kurzinger, K. & Springer, T. J. *Immun.* **127**, 590-595 (1981).
- Dongworth, D. W. & McMichael, A. J. *Br. med. Bull.* **40**, 254-261 (1984).
- Sanchez-Madrid, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7489-7493 (1982).
- Hildreth, J. E. K., Gotch, F. M., Hildreth, D. K. & McMichael, A. J. *Eur. J. Immun.* **13**, 202-208 (1983).
- Lefrançois, L. & Bevan, M. J. *Nature* **314**, 449-452 (1985).
- Haars, R., Rohowsky, C., Reed, E., King, D. W. & Suci-Foca, N. *Immunogenetics* **20**, 397-405 (1984).
- Rohowsky, C., Bonagura, V., Lewison, A., King, D. W. & Suci-Foca, N. *Hum. Immun.* **11**, 173-182 (1984).
- Welte, K. *et al. J. exp. Med.* **160**, 1390-1402 (1984).
- Toyonaga, B., Yanagi, Y., Suci-Foca, N., Miden, M. & Mak, T. W. *Nature* **311**, 385-387 (1984).
- Suci-Foca, N., Rubinstein, P., Popovic, M., Gallo, R. & King, D. W. *Nature* **312**, 275-277 (1984).
- Parham, P. *Immunogenetics* **13**, 509-528 (1981).
- Lampson, L. A. & Levy, R. J. *Immun.* **125**, 293-299 (1980).
- Suci-Foca, N., Rohowsky, C., Kung, P. & King, D. W. *Hum. Immun.* **9**, 37-47 (1984).

Reconstitution of functional receptor for human interleukin-2 in mouse cells

Masanori Hatakeyama, Seijiro Minamoto, Takashi Uchiyama*, Richard R. Hardy, Gen Yamada & Tadatsugu Taniguchi

Institute for Molecular and Cellular Biology, Osaka University, Suita-shi, Osaka 565, Japan

* Department of Internal Medicine, Faculty of Medicine, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Interleukin-2 (IL-2) has a key role in the antigen-specific clonal growth of T lymphocytes, by virtue of its interaction with a specific cell-surface receptor (IL-2R)¹⁻⁴. The growth signal seems to be delivered by IL-2 bound to the high-affinity, but not the low-affinity, receptor^{5,6}. Genes encoding IL-2 (refs 7-13) and its receptor (that is, Tac-antigen)¹⁴⁻¹⁷ have been cloned and analysed in detail. We have now achieved cell-type-specific reconstitution of the high-affinity human IL-2R by expressing the complementary DNA

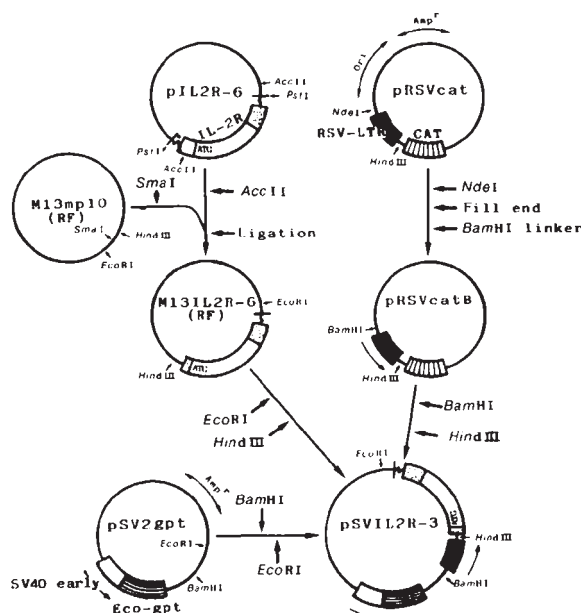


Fig. 1 Construction of the expression plasmid for human IL-2 receptor cDNA. A cDNA library complementary to poly(A)⁺ mRNA of activated human lymphocytes was constructed essentially as described previously¹⁸. Based on the published cDNA sequences for the human IL-2R¹⁴⁻¹⁶, oligonucleotide probes were synthesized and these were used to screen the cDNA library. A clone, pIL2R-6, which contains the entire coding sequence for human IL-2R, was obtained. pIL2R-6 was digested with *AccII* and a 1.55-kilobase (kb) IL-2R cDNA fragment was isolated. This fragment was then ligated to the *SmaI* site of M13 phage mp10 (M13IL2R-G). For the preparation of RSV LTR, pRSVcat²¹ was cleaved with *NdeI* and, after filling-in both ends using DNA polymerase I, a synthetic *BamHI* linker was introduced (pRSVcatB). Then the 1.6-kb *EcoRI*-*HindIII* IL-2R cDNA fragment derived from M13IL2R-6, the 0.6-kb *BamHI*-*HindIII* RSV LTR fragment derived from pRSVcatB and the *EcoRI*-*BamHI*-cleaved pSV2gpt were ligated by T₄ DNA ligase to construct pSVIL2R-3.

cloned from normal lymphocytes. A mouse T-lymphocytic line, EL-4, expressed human IL-2R with high (dissociation constant (K_d) = 160-220 pM) and low (K_d = 2.1-2.2 nM) affinity for recombinant human IL-2, while mouse L929 cells expressed only a single class of the IL-2R with lower affinity (K_d = 34.5 nM) for the ligand. We also show that the human IL-2R expressed in EL-4 cells responds to IL-2 and mediates reversed signal transduction: growth of the EL-4 cells harbouring the IL-2R is inhibited specifically by human recombinant IL-2. The approach described here may provide a general experimental framework for elucidating the molecular basis of signal transduction mediated by specific receptor-ligand interaction.

We prepared a cDNA library using poly(A)⁺ RNA from activated human lymphocytes¹⁸ and screened it for IL-2R (Tac-antigen)-specific clones with oligonucleotide probes. Among several positive clones identified, clone pIL2R-6 was analysed and shown to contain a nucleotide sequence encoding a complete polypeptide identical to that of aberrantly expressed IL-2R in leukaemic cells derived from adult T-cell leukaemia (ATL), reported previously by others¹⁴⁻¹⁶ (data not shown).

Recent studies using radiolabelled IL-2 demonstrated the existence of two classes of IL-2R that differ in their affinities for IL-2 (refs 5, 6). The finding that IL-2-induced T-cell proliferation correlates well with the expression of the high-affinity receptor suggests that the high-affinity receptor alone is capable of signal transduction^{1,2,5}. Although it was reported previously that the cloned human IL-2R cDNA can direct the expression of a cell-surface molecule capable of binding both IL-2 and a monoclonal anti-human IL-2R antibody, anti-Tac^{19,20}, in cultured monkey kidney (COS) cells^{14,15}, it was unclear whether that molecule was functioning as IL-2R, that is, whether it

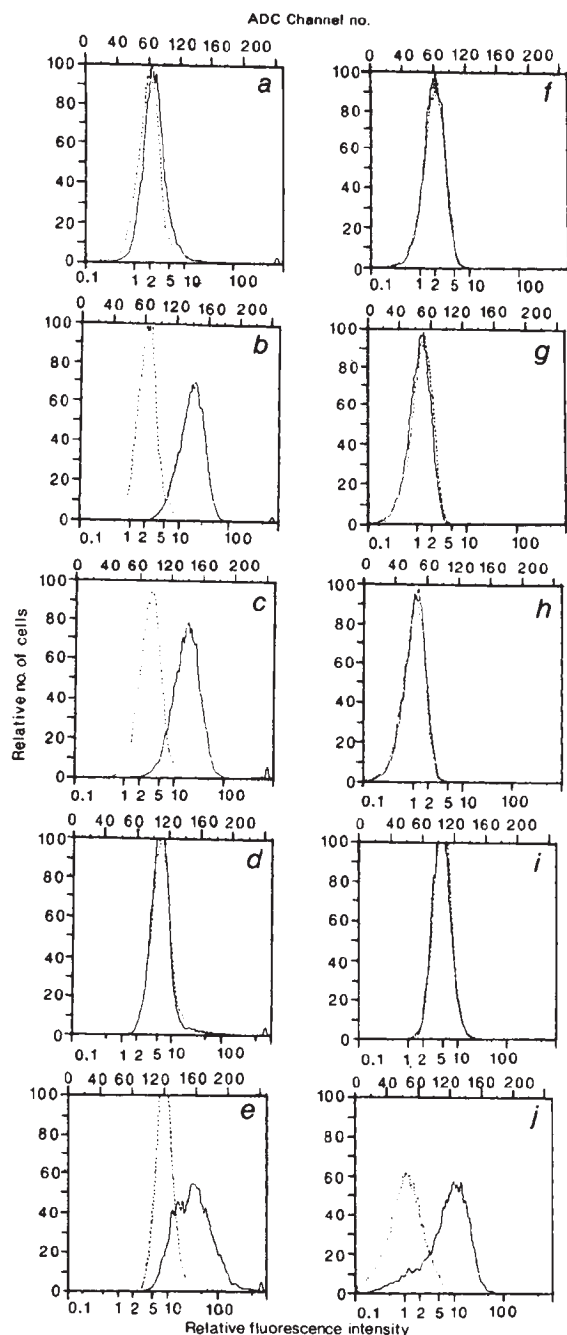


Fig. 2 Flow microfluorimetric analysis of mouse cells expressing either human or mouse IL-2 receptor. Mouse cells were stained indirectly with either mouse anti-human IL-2R monoclonal antibody, anti-Tac (a-e), or rat anti-mouse IL-2R monoclonal antibody, AMT-13 (f-j). Cells used for staining were as follows: a, f, EL-4; b, g, ELT-5 (an EL-4 transformant); c, h, ELT-13 (an EL-4 transformant); d, L929; e, i, LCT-4 (an L929 transformant); j, concanavalin A-activated splenocytes (Con A blasts) of CBN mouse origin. Dotted lines represent staining with the fluorescent second-step conjugate alone.

Methods. Cells (5×10^5) were incubated in 50 μ l Hank's balanced salt solution containing 0.1% bovine serum albumin and 0.1% NaN₃ (HBSS-BN) for 30 min at 4 °C with saturating concentrations of anti-Tac (1:500 dilution of ascites fluid) or AMT-13 (1:500 dilution of ascites fluid). After incubation, cells were washed three times, then resuspended in 40 μ l HBSS-BN containing appropriate concentrations of fluorescein-conjugated goat anti-mouse IgG or mouse anti-rat κ -chain for 30 min at 4 °C. Then the cells were washed three times, resuspended in 0.5 ml HBSS-BN and analysed using a FACS 440 (Becton Dickinson). Con A blasts were prepared as described previously²⁵. EL-4 transformants were produced by transfecting pSVIL2R-3 into EL-4 by the protoplast fusion technique essentially as described by Oi *et al.*³². Mycophenolic acid (4 μ g ml⁻¹)-resistant colonies were visible by 10-14 days; these cells were cloned by the limiting dilution method. Transfection of pSVIL2R-3 into L929 cells was done using the calcium phosphate technique as described by Chu and Sharp³³. After 3 weeks, mycophenolic acid (2.5 μ g ml⁻¹)-resistant cells were collected, stained indirectly with anti-Tac antibody and the positive cells were cloned by single-cell manipulation using FACS, as described by Parks *et al.*³⁴.

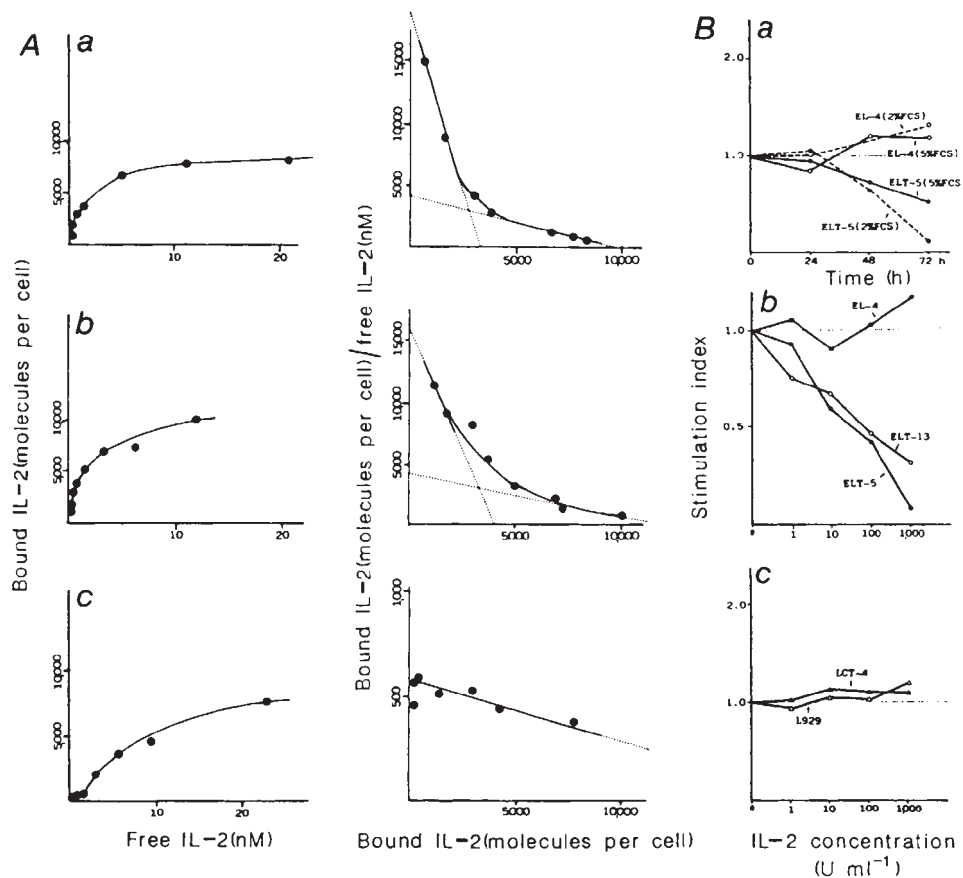
exhibited high affinity for IL-2 and delivered the IL-2-specific signal.

To express the IL-2R cDNA stably in mammalian cells, we first constructed a plasmid, pSVIL2R-3, as shown in Fig. 1. Into this plasmid we inserted the long terminal repeat (LTR) sequence of Rous sarcoma virus (RSV)²¹ just upstream of the IL-2R cDNA and introduced this chimaeric gene into the pSV2gpt vector²². This plasmid was transfected into EL-4 cells, a benzo[a]pyrene-induced thymoma line²³, and into mouse L929 cells, a transformed fibroblast cell line. Transformants were selected in medium containing mycophenolic acid (MPA) and cells expressing the human IL-2R (Tac antigen) were cloned. Blotting analysis of the DNA prepared from each of 16 EL-4 and 10 L-cell transformant clones indicated the presence of one to two copies of the transfected human cDNA (data not shown). Three of these transformed cell clones, ELT-5, ELT-13 (EL-4 origin) and LCT-4 (L929 origin), were arbitrarily chosen for further analysis. As shown in Fig. 2a-e, immunofluorescent staining profiles clearly demonstrate the expression of Tac anti-

gen on the cell surface of IL-2R cDNA transformants, while parental cells are negative for the same antigen.

To determine whether the human IL-2R expressed on the surface of murine cell transformants manifests high or low affinity for IL-2, we performed radiolabelled IL-2 binding assays. Scatchard plot analysis of two EL-4 transformants showed curvilinear features, demonstrating the presence of two distinct receptors having different binding affinities for recombinant IL-2, with dissociation constants of 160 and 2,100 pM in ELT-5, and 220 and 2,200 pM in ELT-13 (Fig. 3A, a, b). This result was highly reproducible and, in our assay system, the K_d of the high-affinity receptor expressed on activated human T lymphocytes and some ATL cells falls within the range 10-100 pM⁶. Thus, the K_d of the high-affinity receptor expressed on the heterologous mouse cell transformants is only slightly lower than that of the endogenous high-affinity receptor on human cells, and the K_d of the second class of receptors expressed on transformant cells is also similar to that of the low-affinity receptor expressed on human cells⁶. In contrast to the EL-4

Fig. 3 **A**, IL-2 binding assay. EL-4 and L929 cell transformants expressing human IL-2R were examined for their IL-2 binding activities using radiolabelled human IL-2. ELT-5 (**a**) and ELT-13 (**b**) were derived from EL-4, and LCT-4 (**c**) was derived from L929. The binding curve of radiolabelled IL-2 (left) and the Scatchard plot analysis of the IL-2 binding data (right) are shown. Human recombinant IL-2 (purity >99.8%) was radiolabelled using ^{125}I -Bolton and Hunter's reagent. The specific radioactivity of the labelled IL-2 was 2,000–5,000 c.p.m. ng^{-1} . The ^{125}I -IL-2 binding assay was performed using methods described by Robb *et al.*⁵, with slight modifications⁶. Serial dilutions of radiolabelled IL-2 and 2×10^6 cells were incubated for 40 min at 37 °C in a total volume of 0.5 ml of RPMI 1640 medium containing 25 mM HEPES, 1% bovine serum albumin (BSA) and 0.1% sodium azide. After the incubation, cells were washed and resuspended in the same medium. The cell suspension (200 μl) in duplicate was layered over 150 μl of a mixture of 20% olive oil and 80% di-*n*-butylphthalate. Following centrifugation of samples, the tips of the sample tubes containing the cell pellets were cut off and the radioactivities were measured using a γ -counter. Nonspecific binding was determined by incubating cells in the presence of a 300-fold excess of unlabelled IL-2. Specific binding was obtained by subtracting nonspecific binding, which represents ~10% of the total binding. **B**, Effect of human recombinant IL-2 on the growth of EL-4 (**a**) and ELT-5 (**b**) in medium supplemented with 2% (—) and 5% (---) fetal calf serum (FCS). **a**, Time course kinetics of the effect of IL-2 (1,000 U ml^{-1}) on the growth of EL-4 (○) and ELT-5 (●) in medium supplemented with 2% (—) and 5% (---) fetal calf serum (FCS). Cells were cultured for 72 h in medium containing 10% FCS in the presence of various concentrations of IL-2. Cells used were EL-4 (■), ELT-5 (●), ELT-13 (○), L929 (△) and LCT-4 (▲). EL-4 and its transformants, ELT-5 and ELT-13, were grown in RPMI 1640 medium supplemented with 10% FCS. L929 and the transformant LCT-4 were cultured in ES medium (Nissui) containing 10% FCS. Cells collected from the culture were plated at 1×10^3 per well in 96-well flat-bottomed microculture plates. Cells were cultured in RPMI 1640 or ES medium in the presence of various concentrations of human recombinant IL-2. Half of the medium was aspirated daily and was replaced with fresh medium containing the same supplements; 4 h (EL-4 and its transformants) or 8 h (L929 and LCT-4) before the termination of culture, 1 μCi of ^3H -thymidine (^3H -TdR) was added to each culture well. Then cells were placed on glass fibre strips using an automatic cell harvester and ^3H -TdR incorporation was measured as an index of cell proliferation. Results are expressed as a stimulation index (SI) = c.p.m. ^3H -TdR incorporation with IL-2/c.p.m. ^3H -TdR incorporation without IL-2. The specific activity of the human recombinant IL-2 was 5×10^7 U per mg protein.



transformants, a mouse L929 transformant, LCT-4, expressed only a low-affinity receptor having a K_d of 34.5 nM (Fig. 3A, c); interestingly, this value is one order of magnitude higher than that of the low-affinity receptor expressed on the EL-4 transformants. The total number of IL-2R molecules per cell was approximately 9,000, 12,000 and 20,000 for ELT-5, ELT-13 and LCT-4, respectively (Fig. 3A). As regards the high-affinity receptor, we found 3,000 molecules per cell in ELT-5 and 4,500 molecules per cell in ELT-13. Thus, the number of IL-2R molecules expressed in these transformants is similar to that of activated human T cells and one order of magnitude lower than that of leukaemic T cells such as MT-1 or HUT-102 (refs 5, 6). As mouse IL-2R is known to react with human IL-2 (ref. 24), we determined whether EL-4 cells and the transformants express mouse IL-2R on their surface. A binding assay using ^{125}I -labelled human recombinant IL-2 demonstrated no significant binding of IL-2 with parental EL-4 cells (data not shown). Moreover, fluorescence-activated cell sorter (FACS) analysis using rat anti-mouse IL-2R monoclonal antibody, AMT-13 (ref. 25), revealed no mouse IL-2R (Fig. 2f–j). These results demonstrate that high-affinity human IL-2R can be reconstituted in heterologous host cells by expression of the cDNA encoding the Tac antigen and that expression of the high-affinity IL-2R requires a T-cell-specific background.

The fact that introduction of human IL-2R cDNA into EL-4

cells leads to constitutive expression of the high-affinity receptor for human IL-2 prompted us to examine whether the human IL-2/receptor system delivers the signal transduction in mouse cells. When recombinant human IL-2 was added to culture media of the transformants, it inhibited the growth of ELT-5 and ELT-13 cells, as expressed by incorporation of ^3H -thymidine into DNA, but had no significant effect on parental EL-4 cells (Fig. 3B, a, b). The growth inhibition of EL-4 transformants was evident at an IL-2 concentration of 1–10 U ml^{-1} (~1.3–13.0 pM), at which concentration IL-2 should bind only to the high-affinity receptor (Fig. 3B, b). Thus, it is likely that the signal transduction is mediated through the high-affinity receptor. At a higher concentration of IL-2 (1,000 U ml^{-1} ; 1.3 nM), a reduction of almost 90% in ^3H -thymidine incorporation was observed. At this IL-2 concentration, growth of the EL-4 transformants was detectably inhibited by 48 h after the onset of culture with IL-2 and ceased almost completely by 72 h (Fig. 3B, a). Cell viability at that time was over 95% and cell proliferation was resumed after removal of IL-2 from the culture media (data not shown). This growth inhibition was not due to growth inhibitory substances whose production is induced by IL-2, as growth of parental EL-4 cells was not inhibited by co-cultivation with ELT-5 or ELT-13 in the presence of 1,000 U ml^{-1} IL-2 (data not shown). In contrast to the observations with these EL-4 transformants, no significant growth alteration was observed for either LCT-4 cells or parental

Table 1 Effect of anti-Tac antibody on the IL-2-induced growth inhibition of ELT-5 cells

Cell line	Treatment		c.p.m. (mean $\pm$ s.d.)	SI
	IL-2	Anti-Tac		
EL-4	-	-	27,600 $\pm$ 1,300	1.00
	500 U ml ⁻¹	-	29,900 $\pm$ 2,200	1.08
	-	1 μ g ml ⁻¹	27,700 $\pm$ 2,400	1.00
	500 U ml ⁻¹	1 μ g ml ⁻¹	29,500 $\pm$ 1,900	1.07
ELT-5	-	-	29,900 $\pm$ 1,300	1.00
	500 U ml ⁻¹	-	5,100 $\pm$ 500	0.17
	-	1 μ g ml ⁻¹	29,700 $\pm$ 4,300	0.99
	500 U ml ⁻¹	1 μ g ml ⁻¹	27,800 $\pm$ 1,300	0.93

The cell proliferation assay was performed as described in the legend to Fig. 3B. Anti-Tac antibody was added at the onset of culture. After 24 and 48 h, half of the medium was removed from each culture and replaced with fresh medium containing the same supplements (IL-2 and/or anti-Tac). After 72 h, 1 μ Ci of ³H-thymidine was added to each culture well, incubation was continued for a further 4 h and the ³H-thymidine incorporation was measured. The stimulation index (SI) was determined as follows; SI = ³H-thymidine incorporation with treatment/³H-thymidine incorporation without treatment.

L929 cells (Fig. 3B, c). As stated above, EL-4 and its transformants do not detectably express mouse IL-2R on their cell surface, therefore growth inhibition of the EL-4 transformants induced by human recombinant IL-2 is most likely mediated by the human IL-2R. This conclusion is reinforced by the finding that the addition of anti-Tac antibody to culture medium containing IL-2 restores almost completely the growth of ELT-5 cells (Table 1). As expected, AMT-13, which recognizes the mouse IL-2R and blocks IL-2-dependent murine T-cell growth²⁵, had no effect on the human IL-2-mediated growth inhibition of ELT-5 (data not shown).

Bifunctional properties of growth factors have been reported also for other growth factor/receptor systems such as that of epidermal growth factor (EGF)^{26,27} and transforming growth factor type β ²⁸. Growth of A431 human epidermal carcinoma cells, which express an unusually high number of EGF receptors, is inhibited in the presence of a high concentration of EGF²⁶. The mechanism of growth inhibition in that system may be related to overstimulation of receptor-associated tyrosine kinase²⁷. As the intracytoplasmic portion of the IL-2R is too short to mediate any kinase activity¹⁴⁻¹⁶, the mechanism of growth inhibition reported here is likely to be different. Recently, high concentrations of IL-2 were reported to inhibit growth of some but not all HTLV (human T-cell lymphotropic virus)-infected T-cell clones, by interacting with IL-2R²⁹. It is illuminating that the growth factor/receptor interaction in neoplastic cells, which are capable of uncontrolled cell proliferation, often leads to reversal of the growth stimulatory signal. As we used a viral LTR sequence for the IL-2R cDNA expression, a 'down-regulatory' mechanism of the heterologous receptor will not operate. Indeed, the level of IL-2R mRNA and Tac-antigen remained unchanged in the presence of IL-2 in ELT-5 cells (our unpublished observations). Thus, continuous growth factor/receptor interaction on the cell surface may also account for the inhibition of the uncontrolled cell growth. Such a reversed signal transduction may be delivered by activation of protein kinase C via a high-affinity receptor³⁰.

The expression of functional, specific receptor molecules introduces a powerful new tool for the investigation of ligand-receptor interactions, and especially for extending our understanding of molecular mechanisms of signal transduction by means of site-directed mutagenesis of the cloned genes and functional study of the gene product.

After submission of this manuscript, Greene *et al.* reported the expression of low-affinity IL-2Rs in mouse L-cells³¹.

We thank Drs J. Hamuro, T. Diamantstein and H. Osawa, B. Howard, P. Berg and H. Matsui for supplying recombinant IL-2, AMT-13, pRSVcat, pSV2gpt and oligonucleotide probes, respec-

tively. We also thank Drs T. Kishimoto, R. Kaempfer, I. Yahara, K. Smith, K. Kumagai, K. Yamasaki and T. Fujita and Ms M. Nagatsuka for help and useful comments. This work was supported in part by a grant-in-aid for Special Project Research, Cancer-Bioscience from the Ministry of Education, Science and Culture of Japan.

Received 11 June; accepted 16 October 1985.

- Smith, K. A. *A. Rev. Immun.* **2**, 319-334 (1984).
- Robb, R. J., Munck, A. & Smith, K. A. *J. exp. Med.* **154**, 1455-1474 (1981).
- Cantrell, D. A. & Smith, K. A. *J. exp. Med.* **158**, 1895-1911 (1983).
- Smith, K. A. & Cantrell, D. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 864-868 (1985).
- Robb, R. J., Greene, W. C. & Rusk, C. M. *J. exp. Med.* **160**, 1126-1146 (1984).
- Uchiyama, T. *et al. J. clin. Invest.* **76**, 446-453 (1985).
- Taniguchi, T. *et al. Nature* **302**, 305-310 (1983).
- Holbrook, N. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1634-1638 (1984).
- Devos, R. *et al. Nucleic Acids Res.* **11**, 4307-4323 (1983).
- Fujita, T., Takaoka, C., Matsui, H. & Taniguchi, T. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7437-7441 (1983).
- Kashima, N. *et al. Nature* **313**, 402-404 (1985).
- Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 68-72 (1985).
- Fuse, A. *et al. Nucleic Acids Res.* **12**, 9323-9331 (1984).
- Leonard, W. J. *et al. Nature* **311**, 626-631 (1984).
- Nikaido, T. *et al. Nature* **311**, 631-635 (1984).
- Cosman, D. *et al. Nature* **312**, 768-771 (1984).
- Shimizu, A. *et al. Nucleic Acids Res.* **13**, 1505-1516 (1985).
- Taniguchi, T., Pang, R. H. L., Yip, Y. K., Henriksen, D. & Vilecek, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3469-3472 (1981).
- Uchiyama, T., Broder, S. & Waldmann, T. A. *J. Immun.* **126**, 1393-1397 (1981).
- Uchiyama, T., Nelson, D. L., Fleischer, T. A. & Waldmann, T. A. *J. Immun.* **126**, 1398-1403 (1981).
- Gorman, C. M., Merlino, G. T., Willingham, M. C., Pastan, I. & Howard, B. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6777-6781 (1982).
- Mulligan, R. C. & Berg, P. *Science* **209**, 1422-1427 (1980).
- Farrar, J. J. *et al. Immun. Rev.* **63**, 129-166 (1982).
- Gillis, S., Ferm, M. M., Ou, W. & Smith, K. A. *J. Immun.* **120**, 2027-2032 (1978).
- Osawa, H. & Diamantstein, T. *J. Immun.* **132**, 2443-2450 (1984).
- Gill, G. N. & Lazar, C. S. *Nature* **293**, 305-307 (1981).
- Buss, J. E., Kudlow, J. E., Lazar, C. S. & Gill, G. N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2574-2578 (1982).
- Roberts, A. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 119-123 (1985).
- Sugamura, K., Nakai, S., Fujii, M. & Hinuma, Y. *J. exp. Med.* **161**, 1243-1248 (1985).
- Farrar, W. L. & Anderson, W. B. *Nature* **315**, 233-235 (1985).
- Greene, W. C. *et al. J. exp. Med.* **162**, 363-368 (1985).
- Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 825-829 (1983).
- Chu, G. & Sharp, P. *Gene* **13**, 197-202 (1981).
- Parks, D. R., Bryan, V. M., Oi, V. T. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1962-1966 (1979).

Human blood platelets showing no response to collagen fail to express surface glycoprotein Ia

H. Karel Nieuwenhuis, Jan Willem N. Akkerman, Wim P. M. Houdijk & Jan J. Sixma

Department of Haematology, University Hospital Utrecht, Catharijnesingel 101, 3511 GV Utrecht, The Netherlands

The interaction of blood platelets with collagen is generally considered to be of primary importance in the arrest of bleeding and to have a role in the pathogenesis of thrombosis and atherosclerosis. Following damage to the vascular endothelium, circulating platelets come into contact with exposed collagen fibrils in the sub-endothelium and spread along it; this is followed by the secretion of several biologically active substances and by aggregation of platelets. The glycoproteins of the platelet plasma membrane have an important role in the mechanisms underlying these processes. So far, two specific defects of platelet function in patients with a bleeding disorder are known to be associated with a glycoprotein defect and the study of these patients has contributed significantly to present concepts of platelet function. The glycoprotein (GP) IIb-III complex, absent or deleted in the aggregation-defective Glanzmann's thrombasthenia¹, has been identified as the platelet fibrinogen receptor. GPIb, which is absent in the adhesion-defective Bernard-Soulier syndrome^{2,3}, has been identified as the von Willebrand factor receptor on platelets. We now report a defect of the platelet plasma membrane glycoprotein composition in a patient whose platelets are totally unresponsive to collagen.

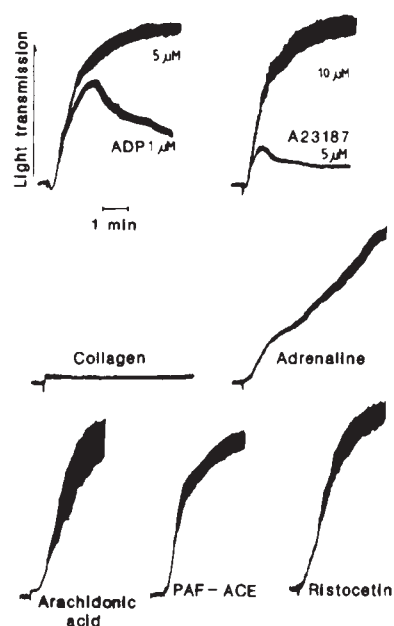


Fig. 1 Aggregation tracings of patient's platelet-rich plasma. Aggregation studies were performed at 37 °C and were normal in response to ADP (1 and 5 μ M), A23187 (5 and 10 μ M), adrenaline (5 μ M), arachidonic acid (1.5 mM), PAF-ACE (1 μ M) and ristocetin (1 mg ml⁻¹). Collagen induced neither shape change nor aggregation.

We studied the platelets of a patient with a haemorrhagic disorder and an excessively long bleeding time (>30 min, compared with the normal range of up to 8 min). Platelet count, coagulation studies and factor VIII parameters were normal. With the informed consent of the patient and control subjects, venous blood was collected in 0.1 vol. of 135 mM trisodium citrate. For some experiments the platelets were isolated by passage through a Sepharose 2B column (Pharmacia) equilibrated in Ca²⁺-free Tyrode's solution (pH 7.2). Aggregation in the patient's platelet-rich plasma (PRP) (Fig. 1) was normal in response to ADP, the ionophore A23187, adrenaline, arachidonic acid, PAF-ACE (platelet aggregating factor/alcohol-chloroform-ether mixture) and ristocetin. Thrombin-induced aggregation was normal in gel-filtered platelet suspensions. In contrast collagen induced neither shape change nor aggregation in PRP and gel-filtered platelet suspensions. This result was found consistently for both low and very high concentrations of several types of fibrillar collagen: equine collagen type I (1–10 μ g ml⁻¹), bovine collagen types I (5–100 μ g ml⁻¹) and III (50–100 μ g ml⁻¹) as well as with purified human collagen types I and III (10–100 μ g ml⁻¹). Normal platelets aggregated with the lowest concentrations of these collagen types. Experiments in which PRP from the patient was mixed with control platelet-free plasma or in which platelet-free plasma from the patient was added to control PRP, demonstrated the absence of a circulating inhibitor and suggested that the defect resided in the platelets.

As current data suggest that the breakdown of phosphatidylinositol (PI) with concurrent formation of diacylglycerol and phosphatidic acid (PA) might be an early common mechanism whereby activation of receptors brings about physiological responses, we studied the metabolism of PI in the patient's platelets (Table 1). In normal platelets labelled with ³²P-orthophosphate, stimulation caused an increase in the ³²P content of PA and PI, which suggests that the PA formed recycles to PI. Treatment of normal ³H-arachidonate-labelled platelets with thrombin or collagen caused loss of radioactivity from PI, with a corresponding accumulation of radioactive PA, free arachidonate and its oxygenation products. In the patient's platelets, thrombin (5 U ml⁻¹) induced a normal decrease of ³H-PI, a normal accumulation of ³²P-PA, a normal incorporation

Table 1 Changes in distribution of label in platelets prelabelled with ³²P-orthophosphate and ³H-arachidonate after stimulation (5 min, 37 °C) with thrombin (5 U ml⁻¹) or collagen (10 μ g ml⁻¹)

	Controls	Patient	
Thrombin			
³ H-PI	33 ± 7%	44%	NS
³² P-PI	146 ± 21%	155%	NS
³² P-PA			
(before exposure)	3,200 c.p.m. per 10 ⁹ pl.	2,100 c.p.m. per 10 ⁹ pl.	
(after exposure)	76,100 c.p.m. per 10 ⁹ pl.	44,300 c.p.m. per 10 ⁹ pl.	
Collagen			
³ H-PI	59 ± 9%	98%	<i>P</i> < 0.05*
³² P-PI	133 ± 15%	95%	<i>P</i> < 0.05*
³² P-PA			
(before exposure)	3,200 c.p.m. per 10 ⁹ pl.	2,100 c.p.m. per 10 ⁹ pl.	
(after exposure)	32,400 c.p.m. per 10 ⁹ pl.	2,000 c.p.m. per 10 ⁹ pl.	

Data are expressed as a percentage (mean ± s.d.) of the radioactivity before activation for ³H-PI and ³²P-PI (for controls, *n* = 5) and, because of low basal values, are given as c.p.m. per 10⁹ platelets (pl.) for ³²P-PA. The ³²P-PA control data are representative of 10 controls studied. Platelets were incubated with 0.01 μ M ³H-arachidonate (specific activity 90 Ci mmol⁻¹; TRK, Amersham International) and 0.1 mCi (carrier-free) H₃³²PO₄ (NEN) for 60 min at 37 °C. Twenty per cent of the radiolabelled arachidonate was taken up by the platelets. After gel filtration the labelled platelets were incubated with thrombin or collagen for 5 min. Extracts were obtained and the phospholipids, fatty acids and thromboxanes were separated by TLC as described elsewhere¹⁹. The ³H and ³²P content of the fractions was determined by scintillation counting.

* Significant difference by Student's *t*-test; NS, not significant (*P* > 0.05).

of ³²P in PI and a normal liberation of ³H-arachidonic acid. Changes in the PI cycle and liberation of ³H-arachidonic acid were not observed after exposure of the patient's platelets to collagen. Thromboxane B₂ (TXB₂) was produced normally in the patient's PRP in response to exogenous arachidonate (1.5 mM), as determined by radioimmunoassay. After incubation for 1 min, the patient's PRP generated 540 nmol TXB₂ per 10¹¹ platelets (normal range 110–518 nmol per 10¹¹ platelets, *n* = 8). There was no detectable TXB₂ generation on stimulation of the patient's platelets with collagen for 5 min (patient < 2, normal range 90–142 nmol per 10¹¹ platelets, *n* = 8).

The interaction of blood platelets with purified human fibrillar collagen type I in a static system was studied according to Legrand⁴ using a technique based on gel filtration of non-adhesive platelets through a Sepharose 2B column after incubation of PRP with fibrillar collagen. For the slightly modified test used here, 0.8 ml of PRP (4 × 10⁵ platelets μ l⁻¹), obtained from EDTA-blood to prevent aggregation, was incubated (37 °C, 2.5 min, without stirring) with 0.2 ml of 0.02 M sodium phosphate buffer pH 7.4 with and without 100 μ g of purified human collagen type I, made fibrillar as described elsewhere⁵. The mixtures were applied to a Sepharose 2B column, equilibrated in a Ca²⁺-free Tyrode's solution, pH 7.4, containing 20 mM HEPES and 0.2% (w/v) bovine serum albumin. The platelets were counted before and after passage and the percentage adhesion calculated. After incubation of the patient's platelets with collagen and subsequent gel filtration, almost no platelets were retained with the collagen fibrils at the top of the column: there was 4.4 ± 0.7% adhesion in the patient, 83.4 ± 1.0% in the control (mean ± s.d. of triplicate experiments). This severe defect in adhesion is present in a static system, where platelet transport towards collagen occurs as a result of diffusion and gravitation, and consequently it differs from the shear-dependent adhesion defects in von Willebrand's disease and Bernard-Soulier syndrome.

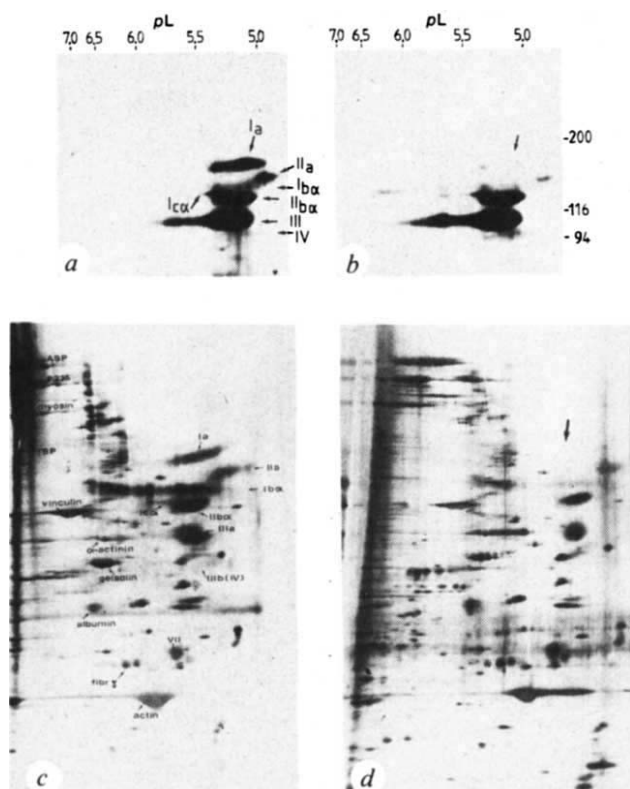


Fig. 2 Autoradiograms (*a, b*) and silver staining (*c, d*) of two-dimensional reduced SDS-polyacrylamide gels for control (*a, c*) and patient (*b, d*) platelets. The positions of the major membrane glycoproteins are indicated. Note that GPIa is lacking in the patient. Values shown on the right-hand side in *b* are $M_r \times 10^{-3}$. **Methods.** Platelets were radioiodinated in parallel by using Iodogen (Pierce Chemicals, Rockford, Illinois) as a catalyst. Isoelectric focusing, SDS-polyacrylamide gel electrophoresis (3–30% acrylamide linear gradient gels) and autoradiography of the gels were performed as described elsewhere²⁰. Gels were silver-stained using a technique described previously²¹.

Analysis of the platelet ectoproteins by silver staining and ¹²⁵I surface labelling, coupled with one- and two-dimensional gel electrophoresis, disclosed a deficiency of the membrane glycoprotein Ia (relative molecular mass (M_r) 168,000 under reducing conditions, pI 4.8–5.4) (Fig. 2). Similar results were obtained when repeated labellings of the patient's platelets were compared with the autoradiograms obtained from 10 controls. Extreme overexposure of the autoradiograms revealed a small amount of residual radioactivity in the GPIa region. To quantify the amount of label in GPIa, this glycoprotein was marked in the gels by autoradiography, cut out of the two-dimensional gels, and counted in a scintillation counter. Several other glycoproteins in the gels were used as standards. Consistent results were obtained on radiolabelling of normal platelets (GPIIb/GPIII = 0.35 ± 0.04 ($\pm$ s.d.); GPIa/GPIII = 0.22 ± 0.06 , $n = 5$). Repeated simultaneous labelling studies of the patient's platelets revealed 15–25% of the normal level of GPIa. Low but detectable amounts of GPIIb (7–30%) and GPIIb-III (12–25%) have been found in some patients with Bernard-Soulier syndrome^{3,6,7} and Glanzmann's thrombasthenia^{8,9}. These patients, referred to as type 2, suffer from a bleeding syndrome and have abnormal (absent) aggregation responses. The defect in the patient studied here—a complete lack of response to collagen associated with a markedly reduced concentration of GPIa—resembles these corresponding platelet function defects in this respect.

GPIa is an ectoprotein which contains intramolecular disulphide bonds as it has an increased apparent relative molecular mass after disulphide cleavage. GPIa is weakly stained by the periodic acid-Schiff reaction, weakly radiolabelled by the per-

iodate method and is hardly affected by treatment with neuraminidase, which indicates that it contains a relatively low concentration of sialic acid. GPIa binds to the lectin wheat-germ agglutinin (*N*-acetylglucosamine specificity)¹⁰.

A putative role as the platelet receptor for collagen has been proposed for glucosyltransferase¹¹, fibronectin¹², complement component C1s¹³, a protein of M_r 65,000 (ref. 14), and glycoprotein IIb¹⁵. Subsequent studies, however, do not confirm such a role for glucosyltransferase^{16,17} or for fibronectin¹⁸ and it is unlikely that GPIIb serves as a collagen receptor, as in Glanzmann's thrombasthenia (a disorder with a deficiency of the fibrinogen receptor, GPIIb-GPIII), collagen-induced shape change and release are normal. In the case of the patient studied here it is evident that the primary interaction of the platelet with collagen is disturbed. In contrast to collagen, several other agonists induced normal physiological responses in the patient's platelets. These results indicate that after exposure to collagen, information could not be transmitted across the plasma membrane to trigger the intracellular events. The observed membrane abnormality occurs in parallel with the lack of responsiveness to collagen, which suggests a cause and effect relationship. Consequently, GPIa may have a receptor or transmitter function.

We thank Marion Schiphorst, Gertie Gorter and Jeanne Korteweg for technical assistance. This work was supported by the Netherlands Organization for the Advancement of Pure Research (ZWO/Fungo) under grant 13-30-92.

Received 3 October; accepted 16 October 1985.

1. Nurden, A. T. & Caen, J. P. *Br. J. Haemat.* **28**, 253–260 (1974).
2. Nurden, A. T. & Caen, J. P. *Nature* **255**, 720–722 (1975).
3. Clemetson, K. J., McGregor, J. L., James, E., Dechavanne, M. & Lüscher, E. F. *J. clin. Invest.* **70**, 304–311 (1982).
4. Legrand, Y. J., Fauvel, E., Kartalis, G., Wautier, J. L. & Caen, J. P. *J. Lab. clin. Med.* **94**, 438–446 (1979).
5. Houdijk, W. P. M., Sakariassen, K. S., Nievelestein, P. F. E. M. & Sixma, J. J. *J. clin. Invest.* **75**, 531–540 (1985).
6. Berndt, M. C., Gregory, C., Chang, B. H., Zola, H. & Castaldi, P. A. *Blood* **62**, 800–807 (1983).
7. Nurden, A. T., Didry, D. & Rosa, J. P. *Blood Cells* **9**, 333–358 (1983).
8. Hagen, I., Nurden, A., Bjerrum, O. J., Solum, N. O. & Caen, J. P. *J. clin. Invest.* **65**, 722–731 (1980).
9. Holahan, J. R. & White, G. C. *Blood* **57**, 174–181 (1981).
10. Clemetson, K. J., Pfeuffer, S. L., Lüscher, E. F. & Jenkins, C. S. P. *Biochim. biophys. Acta* **464**, 493–508 (1977).
11. Jamieson, G. A., Urban, C. L. & Barber, A. J. *Nature new Biol.* **234**, 5–7 (1971).
12. Bensusan, H. B., Koh, T. L., Henry, K. G., Murray, B. A. & Culp, L. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5864–5868 (1978).
13. Tiffany, M. L. & Penner, J. A. *J. Lab. clin. Med.* **96**, 796–802 (1980).
14. Chiang, T. M. & Kang, A. H. *J. biol. Chem.* **257**, 7581–7586 (1982).
15. Tsunehisa, S., Tsuji, T., Tohyama, H. & Osawa, T. *Biochim. biophys. Acta* **797**, 10–19 (1984).
16. Menashi, S., Harwood, R. & Grant, N. E. *Nature* **264**, 670–672 (1976).
17. Santoro, S. A. & Cunningham, L. W. *J. clin. Invest.* **60**, 1054–1060 (1977).
18. Santoro, S. A. & Cunningham, L. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2644–2648 (1979).
19. Holmsen, H., Dangelmaier, C. A. & Holmsen, H. K. *J. biol. Chem.* **256**, 9393–9396 (1981).
20. Sixma, J. J. & Schiphorst, M. E. *Biochim. biophys. Acta* **603**, 70–83 (1980).
21. Morrissey, J. H. *Analyt. Biochem.* **117**, 307–310 (1981).

Specific growth response of *ras*-transformed embryo fibroblasts to tumour promoters

Gian Paolo Dotto, Luis F. Parada & Robert A. Weinberg

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

Chemical carcinogenesis is a process involving multiple steps, as shown in several *in vivo* experimental systems¹. Two early steps have been well characterized: initiation, achieved by a single, subthreshold dose of a carcinogen, and promotion, induced by repetitive treatments with a non-carcinogenic tumour promoter. At the cellular level, establishment of the transformed phenotype is also a multi-step process and activation of several, independent genes appears to be required^{2–4}. Here we show that, like initiated cells, primary rat embryo fibroblasts (REFs) containing a *ras* but

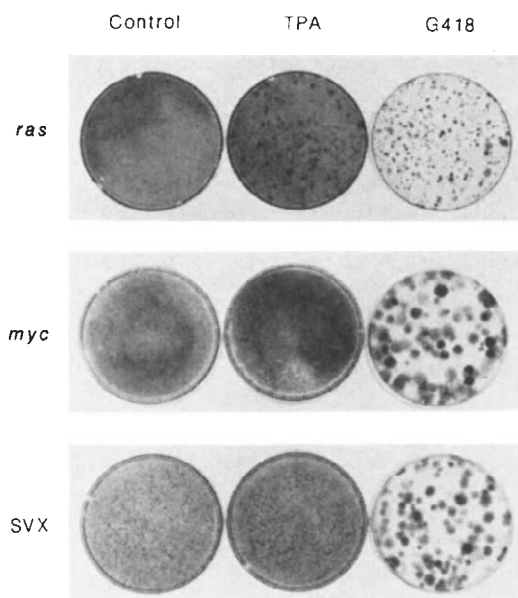


Fig. 1 Effect of TPA on focus formation of oncogene-bearing REFs in monolayer. Tertiary rat embryo fibroblasts (7×10^5 cells per 10-cm dish) were infected with the SVX, raszip 6 or VM viruses (MOI = 0.1–1). After 48 h, cells were split 1:10 into normal medium (Dulbecco's modified essential medium/10% fetal bovine serum), medium plus TPA (10 ng ml^{-1}), or medium plus G418 (0.5 mg ml^{-1}). Dishes were fixed and stained 10 days later.

not a *myc* oncogene, are strongly and specifically stimulated to grow by tumour promoters. In the presence of these promoters, *ras*-containing REFs acquire the ability to overgrow normal cells in the monolayer and to form foci with 100% efficiency. Similar to the *in vivo* situation, promoter effects can be blocked by the concomitant application of retinoic acid.

One of the best *in vivo* models for chemical carcinogenesis is represented by the mouse skin system^{1,5,6}, in which a minimum of two genetic lesions seems to be involved in tumour formation⁷. An apparent parallel arises from the observation that full transformation of primary embryo fibroblasts in culture occurs after introduction of two genetic elements—a *ras*-like and a *myc*-like oncogene. Cells bearing the two oncogenes are fully tumorigenic *in vivo*.

The mouse skin model of initiation and promotion carries specific implications for the action of the promoter and its effects on initiated cells⁵. According to this model, the initiating carcinogen creates a small number of initiated cells, each one bearing a lesion in a specific gene. The promoter then acts to induce the clonal expansion of each initiated cell, which results in a large descendant population, in one cell of which a rare second genetic lesion may now occur. On this basis, an initiated cell is one which enjoys a special growth advantage in the presence of a promoter; and a promoter is a compound that selectively stimulates the outgrowth of already initiated cells.

We reasoned that activation of a *ras*-like or *myc*-like oncogene might represent the initiating event. As a model in which to test this hypothesis, we produced REFs carrying either a *ras* or a *myc* oncogene. We aimed to measure whether either of these cell types enjoyed a special growth advantage in the presence of a tumour promoter such as TPA (12-*O*-tetradecanoylphorbol-13-acetate).

Cells carrying recently introduced oncogenes are often elusive in REF cultures, as their presence is not always signalled by the outgrowth of a focus of transformants². Therefore, we additionally marked these cells with the gene which confers resistance to killing by the antibiotic G418 (ref. 8). This would allow us to quantitate the number of oncogene-bearing cells in a culture at any point by measuring the number that were resistant to the antibiotic (G418^r).

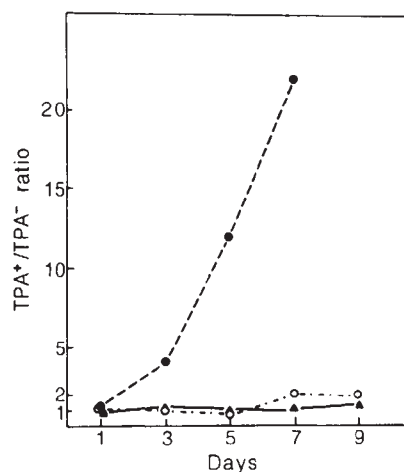


Fig. 2 Effect of TPA on the growth of oncogene-bearing REFs in monolayer. REFs were infected as described in Fig. 1 legend. After 48 h the cells were split 1:10 into either normal or TPA-containing medium and separate dishes were withdrawn at various times after that. Cells were trypsinized, counted and re-plated at a density of 1 and 5×10^5 cells per 10-cm dish in G418-containing medium (0.5 mg ml^{-1} ; Gibco). Colonies were counted after 2 weeks and the ratio of the number of colonies from cells grown with TPA to the number from cells grown without TPA was calculated. Two independent dishes were scored for each experimental point. Values are given for cells infected with raszip 6 (●), VM (○) and SVX (▲).

In previous work with REFs, DNA transfection had been used to introduce oncogenes². However, the efficiency of gene transfer by this method is low, and the number of transfected cells hard to control. Therefore, we used retrovirus vectors to convey both oncogenes and the G418^r marker into these cells. By transfecting the resulting viral genomes into $\psi 2$ cells⁹, we obtained high-titre recombinant viruses that were free of helper. These vectors could introduce oncogenes into REFs but would not allow virus spread beyond the initially infected cell.

One retrovirus carrying the *ras* oncogene was constructed by inserting the viral Ha-*ras* gene (map positions 4,000–4,710; ref. 10) into the *Bam*HI site of SVX, a retroviral vector already containing a G418-resistance gene¹¹. The DNA construct was transfected into $\psi 2$ cells⁹, resulting in a virus stock (raszip 6) showing a high titre of induction of G418 resistance (10^6 – 10^7 colony-forming units ml^{-1}). When the virus was tested on established NIH 3T3 mouse cells, acquisition of resistance to G418 was always accompanied by a morphological transformation characteristic of that induced by the *ras* oncogene. A second virus used in these studies, VM, carries the *myc* oncogene (p110 *gag*-*myc* fusion gene from pv-*myc*²) in addition to the G418^r marker. This virus induces expression of *myc* transcript and protein and is able to immortalize REFs in preparation). As a control, we used the SVX virus¹¹, which carries only the G418 resistance marker and no oncogene.

Third-passage REFs were infected with the three viruses at various concentrations and after 2 days were split into two groups which were then grown in either normal medium or medium containing G418 (0.5 mg ml^{-1}). After incubation for 1 week, dishes were scored either for foci or for colony formation in selective medium.

No foci could be detected in monolayers carrying the *myc* virus, independent of the multiplicity of infection (MOI). Similarly, when the *ras* virus was applied at moderate MOI, resulting in the infection of only 0.1–1% of the cells, no foci were seen. This result agrees with earlier work in which transfection was used to introduce the *ras* oncogene into a small proportion of the REF cells, yielding no observable foci².

We then assessed the effects of TPA on these various populations of oncogene-bearing and normal cells. TPA was used at

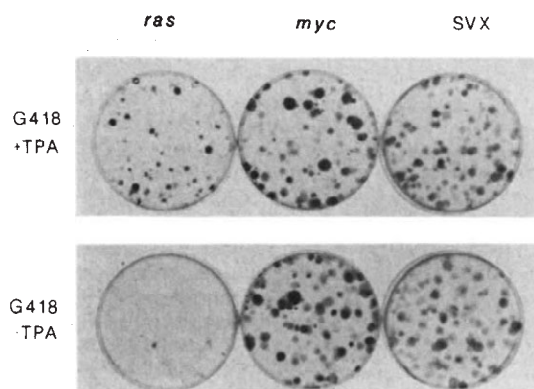


Fig. 3 Effect of TPA on the growth of oncogene-bearing REFs in the absence of surrounding normal cells. Cells were infected as described in Fig. 1 legend and after 48 h were split 1:10 into G418-containing medium (0.5 mg ml^{-1}), with or without TPA (10 ng ml^{-1}). Dishes were stained 1 week later.

10 ng ml^{-1} , a concentration comparable to that normally used to study the physiological effects of this substance on cells in culture¹².

TPA had no effects on the ability of SVX or *myc*-bearing cells to form foci in the monolayer (Fig. 1). In contrast, the effects on *ras*-bearing cells were dramatic, in that large numbers of foci of morphologically transformed cells were now apparent in the monolayer (Fig. 1). Significantly, the number of foci closely approximated the number of G418^r colonies obtained from the same cell population when plated on selective medium (Fig. 1). Most of the cells in these colonies showed a transformed morphology (data not shown). Thus, almost every cell that expressed the G418^r gene of the rasip 6 chimera was also able to form a focus in the presence of TPA. This ability of TPA to affect all the '*ras* cells' present in the monolayer conforms to epigenetic effects associated with promoters and contrasts with the rarely occurring genetic alterations associated with initiating agents¹.

The appearance of these foci suggested that the tumour promoter allowed the *ras* transformants to overcome the growth inhibitory effect imposed on them by the surrounding monolayer of normal cells. We wished to substantiate this view by direct measurement of the number of *ras* (or *myc*)-transformed cells growing in these culture conditions. Cultures infected with SVX or *ras* or *myc* virus were grown for various times in the presence or absence of TPA, then trypsinized and part of them re-plated in the presence of G418, but with no TPA in the medium. This allowed us to measure the proportion of SVX, *ras* or *myc*-bearing cells in the population, independently of any assay of focus-forming ability. It is clear from Fig. 2 that the proportion of '*myc*-cells' was not substantially affected by the presence of TPA. Similarly, the proportion of control SVX-infected cells was also not changed by TPA. In contrast, the proportion of '*ras*-cells' in the cell population was much greater in the presence of the tumour promoter; it was 4-fold higher than the control 3 days after the beginning of the experiment, 10-fold higher at 5 days and 20-fold higher at 7 days. This result demonstrates directly that the clonal expansion of cells containing a *ras*, but not a *myc* oncogene, is strongly and preferentially stimulated by the presence of tumour promoter in the medium.

These data suggest that TPA might act directly on the *ras* transformants, stimulating their growth and thereby allowing them to overcome the inhibitory effects of adjacent normal cells. Alternatively, the promoter might act indirectly, affecting the normal cells and reducing their ability to inhibit growth of the *ras* transformants. To distinguish between these possibilities, we infected cells with the *ras* or the *myc* virus and split them 2 days later into two groups, one being transferred to medium with G418 alone, the other to medium containing G418 plus TPA. In this way, all adjacent normal cells were killed off, and the oncogene-bearing cells could grow up into isolated colonies.

As is clear from Fig. 3, within 1 week after plating, there was a strong effect of TPA on the size, but not the number, of '*ras* cell' colonies in a culture dish. When cells were trypsinized and counted, 10 times more cells were found in the presence of TPA than in its absence. No such difference could be detected in either the control cells or the *myc*-bearing cells. We conclude that at least part of the synergism between TPA and the *ras* oncogene derives from direct effects on the *ras*-bearing cells. While TPA might be able to affect adjacent normal cells, such effects do not need to be invoked to explain our results.

Initiated cells *in vivo* are induced to grow out into papillomas by the continued application of TPA^{1,5,6} and this effect can be blocked by the concomitant application of retinoic acid¹³. As Table 1 shows, retinoic acid exerts an analogous activity in our system, in that it can effectively block the TPA-dependent outgrowth of '*ras*' foci. Thus, TPA and retinoic acid seem to interact similarly *in vivo*, in the promotion and repression of papilloma formation, and *in vitro*, in the stimulation and suppression of growth of *ras*-transformed cells.

Promotion of papillomas can be resolved into two steps^{14,15}. First-stage promotion can be brought about by a single application of TPA, while second-stage promotion requires prolonged applications of either TPA or 'second-stage promoters', such as 12-*O*-retinoylphorbol-13-acetate (RPA) or mezerein, which are either not active or very weakly active in the first stage^{14,15}. These compounds, when tested in our system, were found to be about

Table 1 Effects of various chemicals on focus formation of REFs bearing *ras* oncogene

Chemical	Number of foci per dish
—	0
TPA (10 ng ml^{-1})	55
TPA + RA ($10 \mu\text{M}$)	0
TPA + RA ($1 \mu\text{M}$)	0
TPA + RA ($0.1 \mu\text{M}$)	5
TPA + RA ($0.01 \mu\text{M}$)	45
RPA (100 ng ml^{-1})	60
RPA (10 ng ml^{-1})	40
Mezerein (100 ng ml^{-1})	45
Mezerein (10 ng ml^{-1})	10
4 α -PDD (100 ng ml^{-1})	0
4 α -PDD (10 ng ml^{-1})	0

REFs were infected as described in Fig. 1 legend and after 48 h re-plated (1:10 split) into either normal medium or medium containing the various chemicals; 10 days later, the dishes were stained and foci counted. Two independent dishes were counted for each value shown. TPA and retinoic acid (RA) were purchased from Sigma; RPA, mezerein and 4 α -PDD from LC Services Corp., Woburn, Massachusetts.

as active as TPA (Table 1); this suggests that introduction of an activated *ras* oncogene might bypass the requirements for first-stage promotion and allow these cells to respond directly to second-stage promoters. Another phorbol ester derivative, 4 α -phorbol-12,13-didecanoate (4 α -PDD), which has been shown previously to be inactive both *in vivo* and *in vitro*¹⁶, was also tested in our system and found to have no activity (Table 1).

Activated *ras* and *myc* oncogenes have been found in a large variety of tumours, both of spontaneous origin and chemically induced, and it is assumed that activation of these genes has a causal role in tumour development¹⁷. The present work provides a direct and quantitative demonstration that introduction of an activated *ras*—but not *myc*—oncogene into primary cells causes these cells to respond to a promoter in the manner of an initiated cell. This provides an *in vitro* parallel with work from other laboratories which suggests that *ras* activation occurs at very early times during initiation/promotion of carcinogenesis *in vivo* and may even coincide with the initiating event¹⁸⁻²¹.

Others have recently reported that TPA is able to act synergistically with *myc*²². This observation is only in apparent contrast to ours as these workers used TPA concentrations that were 1,000-fold higher than those used in the present study and they

detected foci of *myc* transfectants in immortalized cells in the presence of TPA only after 6 weeks, in contrast to the 1-week lag period reported here.

TPA, like the *myc* oncogene, is able to collaborate with a *ras* oncogene to allow focal overgrowth in monolayer culture. In this limited sense, TPA acts analogously to *myc*, achieving epigenetically the same effects as those of a genetically activated *myc* gene. We note that TPA has been found to induce expression of the normal cellular *myc* gene²³. This induction may represent the mechanistic basis by which TPA is able to mimic *myc* in the present experimental model.

We thank Midori Maruyama for technical help and Bernard Mathey-Prevot for critical reading of the manuscript. This project was supported by grant CA 39963 from the US NCI and by a grant from the American Business Cancer Research Fund. G.P.D. is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research, which has also supported this investigation with a grant.

Received 14 May; accepted 30 September 1985.

- Hecker, H., Fusenig, N. E., Kunz, W., Marks, F. & Thielman, H. W. (eds) *Carcinogenesis* Vol. 7 (Raven, New York, 1982).
- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-602 (1983).
- Ruley, E. H. *Nature* **304**, 602-606 (1983).
- Newbold, R. F. & Overell, R. W. *Nature* **304**, 648-651 (1983).
- Berenblum, I. in *Cancer: A Comprehensive Treatise* Vol. 1 (ed. Becker, F. F.) 323-344 (Plenum, New York, 1975).
- Boutwell, R. K. *Prog. exp. Tumor Res.* **4**, 207-250 (1964).
- Hennings, H. *et al. Nature* **304**, 67-69 (1983).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
- Mann, R., Mulligan, R. C. & Baltimore, D. *Cell* **33**, 153-159 (1983).
- Dhar, R. *et al. Science* **217**, 934-937 (1982).
- Cepko, C. L., Roberts, B. E. & Mulligan, R. C. *Cell* **37**, 1053-1062 (1984).
- Nishizuka, Y. *Nature* **308**, 693-698 (1984).
- Verma, A. K., Rice, H. M., Shapos, B. G. & Boutwell, R. K. *Cancer Res.* **38**, 793-801 (1978).
- Slaga, T. J., Fischer, S. M., Nelson, K. & Gleason, G. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3659-3663 (1980).
- Furstenberger, G., Berry, D. L., Sorg, B. & Marks, F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7722-7726 (1981).
- Castagna, M. *et al. J. biol. Chem.* **257**, 7847-7851 (1982).
- Land, H., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771-778 (1983).
- Yuspa, S. H., Kilkenny, A. E., Stanley, J. & Lucht, U. *Nature* **314**, 459-462 (1985).
- Balmain, A. & Pragnell, I. B. *Nature* **303**, 72-74 (1983).
- Balmain, A., Ramsden, M., Bowden, G. T. & Smith, J. *Nature* **307**, 658-660 (1984).
- Zarbl, H., Sukumar, S., Arthur, A. V., Martin-Zanca, D. & Barbacid, M. *Nature* **315**, 382-385 (1985).
- Connan, G., Rassoulzadegan, M. & Cuzin, F. *Nature* **314**, 277-279 (1985).
- Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603-610 (1983).

C_{μ} -containing transcripts initiate heterogeneously within the *IgH* enhancer region and contain a novel 5'-nontranslatable exon

Gregory G. Lennon & Robert P. Perry

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

Transcriptional competence of the immunoglobulin heavy-chain locus (*IgH*) is established at an early stage of lymphoid cell development, leading to the appearance of RNA components, previously called C_{μ} RNA¹ or sterile- μ RNA², which contain constant-region sequences but lack variable-region sequences. These components are of two types: those which initiate in the *D* region of alleles that have undergone *DJ_H* (diversity-joining region) rearrangement (D_{μ} transcripts) and those which initiate within the *J_H-C_μ* intron (hereafter termed I_{μ} transcripts)^{3,4}. In pre-B and early B cells, D_{μ} and I_{μ} transcripts are nearly as abundant as the messenger RNA encoding μ heavy chain^{2,3,5}. The D_{μ} transcripts are spliced into RNAs containing *D*, *J_H* and C_{μ} sequences, and in some, but not all, cases these RNAs are translated into D_{μ} proteins⁴. To establish whether the I_{μ} transcripts have any translational potential and to elucidate the structure of their promoter region, we have determined their transcription initiation sites and their mode of splicing. As reported here, by using sequence

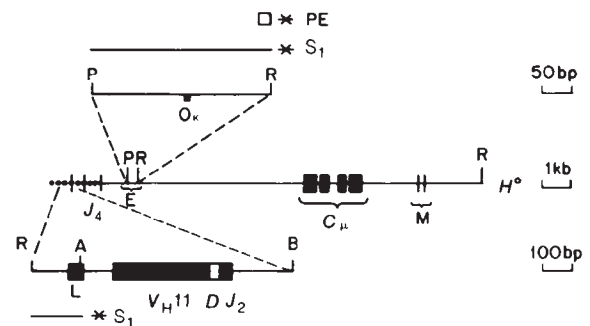
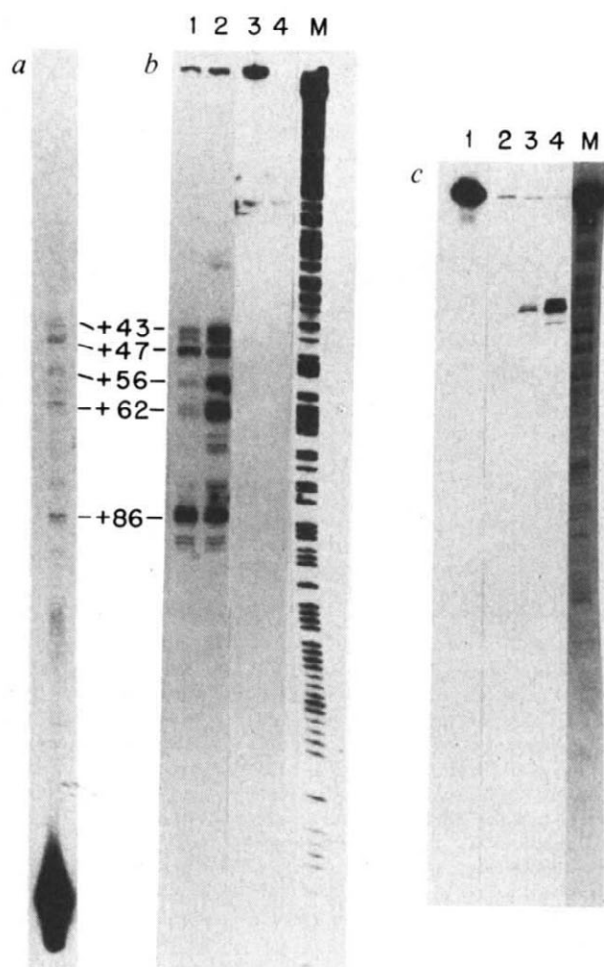


Fig. 1 Schematic diagram of mouse immunoglobulin heavy-chain genes and relevant probes. The germline (H°) region that is common to both alleles of the lymphoma 38C-13 (ref. 3) is shown in the middle of the figure. The upper expansion of the enhancer (E) region shows the location of the octanucleotide (O_x) oriented as in V_{κ} genes. The box to the left of PE (primer extension) indicates the binding position of the 21-mer oligonucleotide. The *PvuII-EcoRI* S_1 probe was isolated from an *Sp6* subclone containing a 700-base pair (bp) *XbaI-EcoRI* insert. The lower expansion shows the 5' region of the productive allele. The *EcoRI-AhaIII* S_1 probe was isolated from subclone p38CV which contains the *EcoRI-BamHI* insert shown. Abbreviations of restriction sites: P, *PvuII*; R, *EcoRI*; B, *BamHI*; A, *AhaIII*. (Not all sites are shown.)

analysis of cloned I_{μ} complementary DNAs, primer extension and S_1 nuclease mapping, we have found that these transcripts have remarkable 5' heterogeneity: there are more than five distinct start sites spanning a region of 44 nucleotides that is located downstream of an octanucleotide found in all variable-region promoters. Such imprecise initiation may result from the lack of a well-defined TATAA motif and the unusual proximity of the octanucleotide to the enhancer region. Approximately 700 nucleotides downstream from these initiation sites, a cryptic splice site is used to create a nontranslatable exon ('nontron') which is joined to the C_{μ} 1 domain. The properties of the nontron may be important for the mechanism of allelic exclusion.

Cloned cDNAs representative of I_{μ} transcripts were obtained from a λ gt10 cDNA library made from RNA of the mouse lymphoma 38C-13. This lymphoma contains cytoplasmic I_{μ} RNA components, presumably transcribed from both productive (H^+) and nonproductive (H^-) alleles, which can be readily distinguished from the mRNA encoding the complete μ heavy chain by their ability to hybridize with a probe specific for a 5' portion of the J_H-C_{μ} intron³. On screening the library with such a probe (IVS₁, ref. 3), we obtained 12 positive clones out of approximately 6×10^5 plaques. The two longest cDNA inserts were purified, subcloned into pUC plasmids, and sequenced. The sequence of their 5' ends mapped within the 0.32-kilobase (kb) *PvuII-EcoRI* fragment previously shown to contain the 5' boundary of some of the C_{μ} RNA components³ as well as the heavy-chain enhancer element^{6,7}.

To localize precisely the sites at which the I_{μ} transcripts initiate, we used both primer extension and S_1 nuclease mapping (Fig. 1). For primer extension of 38C-13 cytoplasmic RNA, we used a 21-mer oligonucleotide representing a sequence near the 3' boundary of the enhancer region. Several products extending 37-80 nucleotides upstream of the primer were observed (Fig. 2a), suggesting the existence of either a multiplicity of transcriptional start sites or a series of reverse transcriptase-dependent artefactual stop sites. S_1 nuclease analysis using a single-stranded *PvuII-EcoRI* fragment allowed us to distinguish between these two possibilities. As a control, we simultaneously determined the start site of the 38C-13 μ mRNA, using the same RNA preparation and an *EcoRI-AhaIII* fragment spanning the leader exon of the productive H^+ gene (Fig. 1). The results of this analysis were clear-cut: the μ mRNA initiates at a unique site (Fig. 2c, lanes 3, 4), whereas the initiation sites of the I_{μ} transcripts are indeed heterogeneous (Fig. 2b, lane 2), corresponding exactly to those seen by the primer extension assay. A similar result was obtained with RNA from C2 cells, a



hybridoma derivative of 38C-13 which contains only the nonproductive (H^-) allele³ (Fig. 2b, lane 1). There appears to be a difference in the relative abundance of the various fragments protected by 38C and C2 RNA, suggesting a possible difference in the relative utilization of the different I_μ start sites in these two cells. Whether this difference is related to a difference in I_μ transcription on the H^+ and H^- alleles remains to be established.

The I_μ start sites are located (Fig. 3a) 43–86 nucleotides downstream of a highly conserved^{8,9} octanucleotide (ATTTCAT) which is found in all V_κ genes and, in inverse orientation, in all V_H genes about 60–80 nucleotides upstream of the cap site. This octanucleotide element is apparently essential for the transcriptional activity of these genes^{9,10}. *In vivo* footprinting assays¹¹ have indicated that almost all tissue-specific protein binding in the enhancer region occurs upstream of this octanucleotide (Fig. 3a), suggesting that this region may contain binding sites for factors that are important for the generation of I_μ transcripts. Comparing the position of the octanucleotide implicated in I_μ transcription with that of an octanucleotide brought upstream of the enhancer by the productive rearrangement of a V_H gene, we note that the I_μ octanucleotide is located on the opposite side of these binding sites and in inverse orientation. Thus, the polarity relative to the binding sites is the same for both I_μ and H^+ octanucleotides.

The 38C H^+ gene promoter has a TATAA box correctly spaced upstream of the unique initiation site (Fig. 3b), whereas for the I_μ transcripts there is no unique, well-defined TATAA box. Initiation heterogeneity similar to that observed for I_μ transcription has been observed at sites adjacent to the simian virus 40 (SV40) and polyoma enhancers^{12,13}. The viral promoters responsible for these transcripts also lack TATAA box homologies. To date, very few cellular genes with this extent of

Fig. 2 Mapping of transcription initiation sites. **a**, Primer extension. 50 ng of 21-mer labelled with [γ -³²P]dATP was hybridized to 10 μ g of 38C cytoplasmic poly(A)⁺ RNA for 33 h at 37 °C in 20 μ l of a solution containing 80% formamide³¹ and 21 μ g of poly(A)⁺ cytoplasmic carrier RNA obtained from MPC 11 myeloma cells. Most of the unannealed primer was removed by oligo(dT)-cellulose chromatography, and the remaining material extended using avian myeloblastosis virus reverse transcriptase. Products were loaded onto a 10% polyacrylamide/7 M urea thickness-gradient gel. Numbering indicates the distance (in nucleotides) downstream from the first nucleotide of the conserved octanucleotide. The strong band at the bottom is non-extended primer. **b**, S_1 protection assay of I_μ transcripts. RNA samples were annealed to a gel-separated minus-strand probe (upper expansion, Fig. 1) for 3 h at 40 °C, then the samples were diluted and treated for 37 min at 19 °C with 4,000 units of S_1 nuclease³². Lane 1, 20 μ g C2 cytoplasmic poly(A)⁺ RNA; lane 2, 20 μ g 38C cytoplasmic poly(A)⁺ RNA; lanes 3, 4, no RNA (S_1 nuclease treatment omitted for lane 3). **c**, S_1 protection assay of H^+ transcripts. Same general protocol as for **b**. The minus-strand probe is shown in the lower extension of Fig. 1. Lanes 1 and 2, 1 and 5 μ g of carrier RNA, respectively (S_1 nuclease treatment omitted for lane 1); lane 3, 1 μ g 38C cytoplasmic poly(A)⁺ RNA; lane 4, 5 μ g 38C cytoplasmic poly(A)⁺ RNA. M, G + A sequencing reactions of the appropriate S_1 probe plus strands, one of the three such sequencing reactions used as markers. Transcription initiation sites were determined from a comparison of S_1 nuclease-resistant fragments and marker lanes, applying the 4.5-nucleotide correction factor established in previous experiments³³.

heterogeneity have been described, although one interesting exception is the hydroxymethyl glutaryl coenzyme A reductase gene, which initiates heterogeneously in a G + C-rich region that lacks a TATAA box¹⁴. However, as several other genes, including even some V_H and V_κ genes, can initiate precisely despite the lack of a good TATAA motif (see refs 15 and 16 for examples), it is clear that the absence of this motif does not necessarily engender transcriptional heterogeneity. Perhaps the unusual proximity of the octanucleotide to the major protein-binding region of the enhancer also contributes to the imprecise initiation of I_μ transcripts.

Knowledge of the cDNA sequences allowed us to identify the 5' splice site used in the processing of the I_μ transcripts. The junction between the J_H - C_μ intron sequence and the C_μ 1 exon is the same in both cDNA sequences; it defines a splice site 735 nucleotides downstream from the most 5' of the transcriptional start sites (Fig. 3c). The sequences flanking this splice site, which is normally not recognized in the splicing of pre- μ mRNA or D_μ transcripts, conform reasonably well to the consensus for 5' splice sites¹⁷. Utilization of this cryptic splice site creates a large exon that contains over 35 termination codons in the three reading frames (Fig. 4). Indeed, in the frame that encodes the C_μ protein, there are more than 10 stop codons distributed over the entire exon segment, and utilization of the first non-blocked initiation codon would eliminate 25 amino acids from the C_μ 1 domain. Moreover, according to present ideas about translation initiation¹⁸, such an exceptionally long nontranslatable 5' exon (nontron), with at least 12 AUG codons, should greatly reduce the efficiency with which I_μ transcripts can be translated into any protein with C_μ determinants. Thus, it is now evident why such proteins are not detected in cells that are demonstrably producing I_μ transcripts¹⁹.

A matrix-plot comparison of the mouse and human J_H - C_μ intron sequences revealed a remarkable (for intron sequences) conservation of the 5' portion of this intron²⁰. The conserved segment not only embraces the enhancer region, including a 7/8 octanucleotide homologue, but in fact extends slightly beyond the 3' border of the nontron and includes its splice junction and over 40 stop codons. The fact that both the cryptic splice site and the multiplicity of stop codons are evolutionarily conserved suggests that the nontron may be of functional significance. What could be the significance of an element that prevents C_μ translation? One possibility which we find attractive is related to mechanistic models of allelic exclusion^{21–23}. An important attribute of these models is a feedback system which causes the

Fig. 3 Sequence characteristics of I_{μ} transcripts. *a*, A 298-nucleotide sequence of the IgH enhancer region (as listed in GenBank and confirmed by us) showing the transcription initiation sites (*). The *PvuII* site corresponds to that shown in Fig. 1. The octanucleotide is boxed; ▼ indicates the nucleotide where the longest cDNA begins; ○ and ●, respectively, indicate protection from, and enhancement of, reactivity with dimethyl sulphate by lymphocyte-specific factors¹¹. *b*, Sequence of the 5' region of the 38C H^+ gene (downstream of the *EcoRI* site, determined by chemical cleavage; upstream, taken from the sequence³³ of a closely related gene, V_H11) showing the initiation site (*), as well as the octanucleotide and TATAA motifs (boxed). Horizontal arrows under the octanucleotides emphasize the inverted polarity of these sequences in I_{μ} and V_H promoters. *c*, Sequence spanning the site (/) at which the 5' portion of the I_{μ} transcript is spliced to the C_{μ} exon, as determined from the cDNA sequence. The sequence of this region agrees with that listed in GenBank.

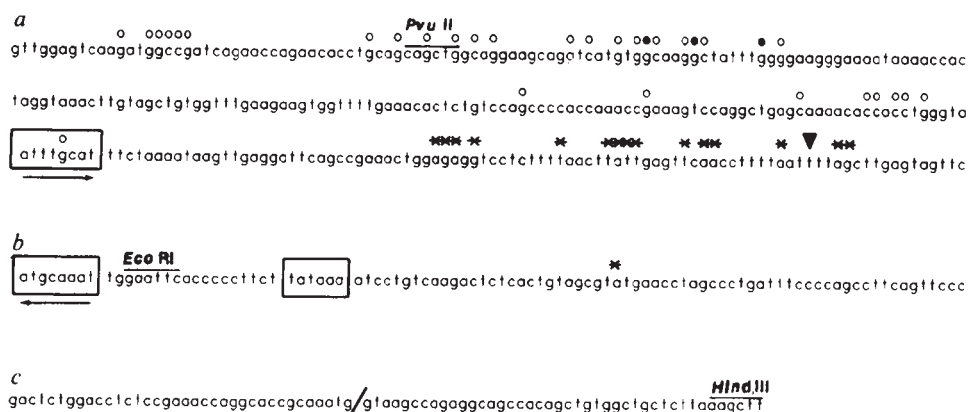
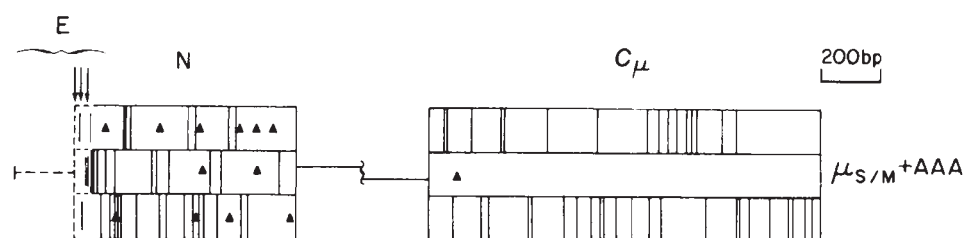


Fig. 4 Schematic diagram of the processed portions of I_{μ} transcripts. N, nontron; E, enhancer region; ↓↓, initiation sites; C_{μ} , entire C_{μ} region, which is joined to either a secretory (μ_S) or a membrane-associated (μ_M) polyadenylated terminus. Reading frames are aligned according to the splice junction shown in Fig. 3c. ▲, Location of AUG codons in the nontron (the first AUG in the C_{μ} coding frame is also indicated); vertical bars represent stop codons.



developing B lymphocyte to cease rearranging its immunoglobulin genes when a productive rearrangement has been made. There are two stages at which the cell apparently recognizes such feedback signals: first, when a heavy-chain gene has been productively rearranged, and later when a productive light-chain rearrangement occurs. While the second signal could be related to the appearance of a completely assembled immunoglobulin on the cell surface, the first would presumably result from the appearance of intracellular μ heavy chain. If the C_{μ} region were the critical determinant for this feedback signal, then any translation products of the I_{μ} transcripts could repress the rearrangement of heavy-chain genes prior to the formation of a productive allele. Thus, we propose that the nontron prevents the generation of a feedback signal which would prematurely curtail heavy-chain gene rearrangement. We are implicitly assuming here that early transcriptional activity at the heavy-chain locus is essential for, or at least an obligatory concomitant of, the rearrangement mechanism. The fact that C_{κ} (refs 24, 25), V_{λ} (ref. 26) and some V_H (ref. 27) regions are transcriptionally active prior to rearrangement lends support to this notion.

If our proposal is correct and any C_{μ} -containing polypeptide can serve as a feedback signal, we must presume that the so-called D_{μ} proteins, which are produced when the 5'-flanking regions of certain D elements are joined in the proper reading frame to a J element, would also inhibit further rearrangement. Although examples of Abelson virus-transformed pre-B-cell lines that produce D_{μ} proteins and are still capable of undergoing VD rearrangement have been reported⁴, more recent findings²⁸ indicate that such rearrangement occurs only when the level of D_{μ} protein is very low, and that high levels of D_{μ} protein may, in fact, be inhibitory for recombination. The I_{μ} RNAs are more abundant than D_{μ} RNA (ref. 3 and our unpublished observations) and are generated over a longer time span since their production is independent of the rearrangement status of the H alleles. Hence, if there were translation products of I_{μ} RNAs, they would probably be at sufficiently high con-

centrations to repress recombination. We also find it noteworthy that for most H^+ loci in which the D segment can be unambiguously identified, the $D-J$ join is in a frame that could not have previously produced a D_{μ} protein^{29,30}. Thus, there may actually be an inverse correlation between the ability to produce D_{μ} protein and the subsequent possession of a completely rearranged H allele.

Irrespective of its implications for allelic exclusion, the properties of the nontron may represent a means of accommodating a possible requirement for transcriptional competence at the J_H-C_{μ} locus while avoiding the deleterious consequences of a protein potentially encoded by the transcription product.

We thank Kathleen Nelson, Leanne Wiedemann and Michal Fried for helpful discussions, Tom St John and Irving Weissman for the 38C cDNA library, Michael Kriegler for the oligonucleotide primer, Mark Shlomchik for primer extension reagents, Dawn Kelley for the p38CV subclone, and Gerald Siu for the unpublished V_H11 sequence. This work was supported by grants from the NSF and NIH to R.P.P., and by an appropriation from the Commonwealth of Pennsylvania. G.G.L. acknowledges support from a NIH predoctoral training grant to the University of Pennsylvania.

Received 22 July, accepted 30 September 1985.

- Kemp, D. J., Harris, A. W. & Adams, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7400-7404 (1980).
- Alt, F. W., Rosenberg, N., Enea, V., Siden, E. & Baltimore, D. *Molec. cell. Biol.* **2**, 386-400 (1982).
- Nelson, K. J., Haimovich, J. & Perry, R. P. *Molec. cell. Biol.* **3**, 1317-1332 (1983).
- Reth, M. G. & Alt, F. W. *Nature* **312**, 418-423 (1984).
- Perry, R. P. & Kelley, D. E. *Cell* **18**, 1333-1339 (1979).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
- Gillies, S. T., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-728 (1983).
- Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2650-2654 (1984).
- Falkner, F. G. & Zachau, H. G. *Nature* **310**, 71-74 (1984).
- Bergman, Y., Rice, D., Grosschedl, R. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7041-7045 (1984).
- Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. *Science* **227**, 134-140 (1985).
- Ghosh, P. K., Reddy, V. B., Swinscoe, J., Lebowitz, P. & Weissman, S. M. *J. molec. Biol.* **126**, 813-846 (1978).

13. Cowie, A., Tyndall, C. & Kamen, R. *Nucleic Acids Res.* **9**, 6305-6322 (1981).
14. Reynolds, G. A. *et al. Cell* **38**, 275-285 (1984).
15. Kelley, D. E., Coleclough, C. & Perry, R. P. *Cell* **29**, 681-689 (1982).
16. Wiedemann, L. M. & Perry, R. P. *Molec. cell. Biol.* **4**, 2518-2528 (1984).
17. Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
18. Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
19. Walker, I. D. & Harris, A. W. *Nature* **288**, 290-293 (1980).
20. Mills, F. C., Fisher, L. M., Kuroda, R., Ford, A. M. & Gould, H. *Nature* **306**, 809-812 (1983).
21. Alt, F. W., Enea, V., Bothwell, A. L. M. & Baltimore, D. *Cell* **21**, 1-12 (1980).
22. Coleclough, C., Perry, R. P., Karjalainen, K. & Weigert, M. *Nature* **290**, 372-378 (1981).
23. Hieter, P. A., Karsmeyer, S. J., Waldman, T. A. & Leder, P. *Nature* **290**, 368-372 (1981).
24. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
25. Nelson, K. J. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5305-5309 (1985).
26. Picard, D. & Schaffner, W. *EMBO J.* **3**, 3031-3035 (1984).
27. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271-281 (1985).
28. Reth, M. G., Ammirati, P., Jackson, S. & Alt, F. W. *Nature* **317**, 353-355 (1985).
29. Coleclough, C. *Trends Immun.* **6**, 128-130 (1985).
30. Kaartinen, M. & Mäkelä, O. *Immun. Today* (in the press).
31. Casey, J. & Davidson, N. *Nucleic Acids Res.* **4**, 1539-1552 (1977).
32. Sharp, P. A., Berk, A. & Berget, S. *Meth. Enzym.* **65**, 750-768 (1980).
33. Wagner, M. & Perry, R. P. *Molec. cell. Biol.* (in the press).

Resistance to β -lactam antibiotics by re-modelling the active site of an *E. coli* penicillin-binding protein

Philip J. Hedge & Brian G. Spratt

Microbial Genetics Group, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK

The β -lactam antibiotics kill bacteria by inhibiting a set of penicillin-binding proteins (PBPs) that catalyse the final stages of peptidoglycan synthesis^{1,2}. In some bacteria the development of intrinsic resistance to β -lactam antibiotics by the reduction in the affinity of PBPs causes serious clinical problems^{1,3-11}. The introduction of β -lactam antibiotics that are resistant to hydrolysis by β -lactamases may also result in the emergence of intrinsic resistance among the Enterobacteriaceae¹. The clinical problems that would arise from the emergence of resistant PBPs in enterobacteria have led us to examine the ease with which *Escherichia coli* can gain resistance to β -lactams by the production of altered PBPs. The development of resistant PBPs also provides an interesting example of enzyme evolution, since it requires a subtle re-modelling of the enzyme active centre so that it retains affinity for its peptide substrate but excludes the structurally analogous^{2,12,13} β -lactam antibiotics. We show here that only four amino-acid substitutions need to be introduced into PBP 3 of *E. coli* to produce a strain possessing substantial levels of resistance to a wide variety of cephalosporins. We also show that transfer of the gene encoding the resistant PBP 3 from the chromosome to a plasmid could result in the spread of intrinsic resistance not only to other strains of *E. coli* but also to other enterobacterial species.

Most β -lactam antibiotics in clinical use kill *E. coli* at their minimum inhibitory concentrations by inactivation of PBP 3, resulting in the inhibition of cell division and the growth of the bacteria into filamentous cells^{1,14}. Many of these antibiotics (for example, cephalixin, cefuroxime, third-generation cephalosporins, mezlocillin, piperacillin and monobactams¹⁵) show much greater affinity for PBP 3 than for the other essential PBPs, and therefore kill *E. coli* over a wide concentration range exclusively by inactivation of PBP 3¹. Substantial levels of resistance to these β -lactams could therefore be achieved by the development of a form of PBP 3 that has reduced affinity for the antibiotics. Amino-acid substitutions that allow PBP 3 to discriminate between the binding of cephalixin and its normal substrate are, however, very rare and result in only a low level of resistance¹⁶. Mutants with high levels of resistance require multiple amino-acid substitutions in PBP 3 which re-model the active centre to reduce its affinity for β -lactams without impairing its ability to process the normal substrate. We have followed

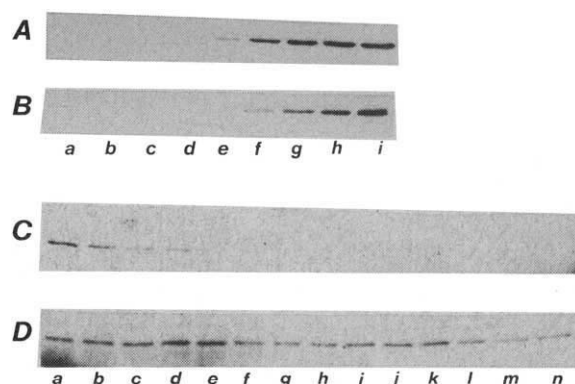


Fig. 1 Affinity of PBP 3 of the fourth-level mutant (SP1124) for benzylpenicillin and cephalixin²⁶. Increasing concentrations of ³H-benzylpenicillin (21 Ci mmol⁻¹; a gift of Merck Sharp and Dohme) were incubated for 10 min at 30 °C with washed cell envelopes prepared from *E. coli* expressing wild-type PBP 3 (strain C600pPH105) (A) or the resistant PBP 3 of the fourth-generation mutant (strain SP1124pPH105pbpB^r) (B). The reactions were terminated and the PBPs fractionated on a 10% SDS polyacrylamide gel as described elsewhere²⁶. The levels of ³H-benzylpenicillin bound to the PBPs were determined by fluorography. Only the section of the fluorograph showing PBP 3 is illustrated. ³H-benzylpenicillin concentrations (μg ml⁻¹): a, 0.06; b, 0.12; c, 0.25; d, 0.5; e, 1; f, 2; g, 4; h, 8; i, 16. A measure of the affinity of the PBPs for cephalixin was obtained by incubating washed cell envelopes of strain C600pPH105 (C) or SP1124pPH105pbpB^r (D) with increasing concentrations of cephalixin for 10 min at 30 °C and then adding ³H-benzylpenicillin (30 μg ml⁻¹ final concentration) for a further 10 min to label the PBPs that are not complexed with cephalixin²⁶. The PBPs were separated and visualized as above. Only the section of the fluorograph showing PBP 3 is illustrated. Cephalixin concentrations (mg ml⁻¹): a, 0; b, 0.015; c, 0.03; d, 0.06; e, 0.12; f, 0.24; g, 0.47; h, 0.95; i, 1.9; j, 3.8; k, 7.5; l, 15; m and n, 30.

the evolution of *E. coli* PBP 3 during selection for increasing levels of resistance to cephalosporins.

The most resistant of the cephalixin-resistant mutants obtained in our previous study¹⁶ (SP1091; Table 1) was mutagenized separately with ethyl methane sulphonate, ultraviolet irradiation and nitroquinoline *N*-oxide, and second-level mutants showing increased levels of resistance to cephalixin were selected at 37 °C. Those mutants in which the increased resistance was due to a further reduction in the affinity of PBP 3 for cephalixin were distinguished from the majority of the mutants in which resistance was due to decreased outer membrane permeability¹⁷ or increased expression of the chromosomal β -lactamase¹⁸, by the phage P1-mediated co-transduction of the resistance with the *leu* genes which map very close to the PBP 3 gene (*pbpB*)¹⁹. However, second-level PBP 3 mutants that showed a substantial increase in resistance could not be obtained from any of the mutagenized cultures and the best mutants obtained showed only a slight increase in cephalixin resistance. The decrease in the affinity of PBP 3 for cephalixin in SP1091 resulted in a decrease in its thermostability so that the strain grew well at 37 °C but filamented and died at 43 °C¹⁶. The second-level mutants had regained the normal thermostability of PBP 3 and the ability to form colonies at 43 °C.

One of the second-level mutants (SP1096, a transductant carrying the mutant *pbpB* gene in an unmutagenized background) was treated with each of the mutagens and selection was applied for a further increase in cephalixin resistance; those third-level mutants in which the increased resistance was due to PBP 3 changes were identified. Repeated attempts to obtain third-level mutants produced only one isolate (SP1097), which showed a further substantial increase in PBP 3-mediated cephalixin resistance. The decrease in affinity of PBP 3 of SP1097 was accompanied by decreased thermostability and the

Table 1 Properties of a series of PBP 3 mutants with increasing levels of resistance to cephalosporins

Strain	Minimum inhibitory concentration ($\mu\text{g ml}^{-1}$) of														Growth at 43 °C
	Clxn	Cdrx	Crdn	Cthn	Crxm	Cmdl	Ctxm	Ctaz	Cprz	Ampn	Pipn	Mzln	Ticn	Aztr	
C600 (parent)	6	10	20	5	5	1	0.1	0.2	0.2	5	2	2	5	0.1	Yes
SP1091 (first level)	50	100	200	100	100	2	0.1	0.1	0.1	1	0.5	0.5	0.5	0.05	No
SP1096 (second level)	60	100	100	20	10	5	0.1	0.1	0.5	10	2	2	1	0.1	Yes
SP1097 (third level)	500	400	1,000	200	200	40	1	1	0.5	5	1	1	1	0.05	No
SP1124 (fourth level)	500	400	2,000	200	500	40	2	2	1	5	5	2	2	0.1	Yes

The isolation of the first-level mutant, SP1091, and the methods for identifying *pbpB* mutants, have been described elsewhere¹⁶. The mutants with increasing levels of resistance to cephalosporins were isolated after mutagenesis with ethyl methane sulphonate²⁰, ultraviolet irradiation²⁰ or nitroquinoline *N*-oxide²¹ as described in the text. Determination of minimal inhibitory concentrations: 10- μl amounts of a 10^{-5} dilution of overnight cultures of the mutants and the parent strain (~ 150 colony-forming units) were spotted on L-agar containing a range of concentrations of the antibiotics. The plates were incubated at 37 °C for 24 h and the minimum inhibitory concentration was the lowest concentration of antibiotic that prevented bacterial growth. A similar number of cells were plated on L-agar at 43 °C to test the ability of the mutants to grow at this temperature. Abbreviations for cephalosporins: Clxn, cephalexin; Cdrx, cefadroxil; Crdn, cephradine; Cthn, cephalothin; Crxm, cefuroxime; Cmdl, cefamandole; Ctxm, cefotaxime; Ctaz, ceftazidime; Cprz, cefoperazone. Abbreviations for penicillins: Ampn, ampicillin; Pipn, piperacillin; Mzln, mezlocillin; Ticn, ticarcillin. Aztr is aztreonam, a monobactam¹⁵.

Table 2 Nucleotide changes in the *pbpB* gene, and the predicted amino-acid substitutions in PBP 3, of a series of mutants showing increasing levels of resistance to cephalosporins

Strain	Mutagen used for isolation	Nucleotide change in <i>pbpB</i>	Amino-acid substitution in PBP 3
SP1091 (first level)	Ethyl methane sulphonate	AAC to AGC	Asn 361 to Ser
SP1096 (second level)	Ethyl methane sulphonate	GTT to ATT	Val 530 to Ile
SP1097 (third level)	Ultraviolet irradiation	TAC to CTC	Tyr 541 to Leu
SP1124 (fourth level)	Ethyl methane sulphonate	GAA to AAA	Glu 349 to Lys

The wild-type *pbpB* gene was cloned into the pSC101-derived vector, pLG339, to produce pPH105²². The mutations in the *pbpB* gene of each of the mutants were transferred into pPH105 *in vivo* by homogenization as described previously¹⁶. The wild-type and mutant *pbpB* genes were cloned from these plasmids into phage M13 mp8/9 vectors²³ and were dideoxy-sequenced²⁴ on both strands using a set of oligonucleotide primers complementary to regions along each strand of the *pbpB* gene as described previously¹⁶. The nucleotide sequence of the wild-type *pbpB* gene has been determined by Nakamura *et al.*²⁵. The nucleotide changes shown are the new mutations that have occurred at each level; in addition, the *pbpB* gene of the mutants contained the mutations they obtained at the earlier levels.

Table 3 Conferral of cephalosporin resistance on enterobacteria by the introduction of a plasmid expressing a cephalosporin-resistant form of PBP 3

Strain	Minimum inhibitory concentrations ($\mu\text{g ml}^{-1}$) of:						
	Clxn	Crdn	Cthn	Crxm	Cmdl	Ctxm	Cprz
<i>E. coli</i> pPH105	6	20	5	5	1	0.1	0.2
<i>E. coli</i> pPH105 <i>pbpB</i> ^r	500	2,000	200	500	50	2	1
<i>S. typhimurium</i> pPH105	5	20	1	2	<0.2	0.1	0.5
<i>S. typhimurium</i> pPH105 <i>pbpB</i> ^r	500	500	20	100	5	0.5	2
<i>K. pneumoniae</i> pPH105	5	10	2	2	0.5	0.05	0.5
<i>K. pneumoniae</i> pPH105 <i>pbpB</i> ^r	500	1,000	50	100	50	0.5	5

pPH105, expressing wild-type PBP 3, and pPH105*pbpB*^r, expressing the resistant PBP 3 of the fourth-level mutant SP1124, were introduced by transformation into *E. coli* C600, *S. typhimurium* LR5000 and *Klebsiella pneumoniae* UNF122. The abbreviations for antibiotics and the method for measuring the minimum inhibitory concentrations are as in Table 1.

strain filamented at 43 °C. The increased resistance of strain SP1097 was due to a double nucleotide change in *pbpB* (Table 2). Presumably SP1097 was obtained at such a low frequency because there were no single-nucleotide changes in the *pbpB* gene of the second-level mutant that produced the desired phenotype.

SP1097 lysed when grown in the presence of its minimum inhibitory concentration of cephalexin, indicating that the affinity of PBP 3 had been reduced to the extent that killing was now occurring by the inactivation of PBP 1A/1B rather than PBP 3. The mutant had therefore achieved the maximum level of resistance to cephalexin that could be achieved by a reduction in the affinity of PBP 3 and had also gained close to the maximum possible levels of cross-resistance to a number of other cephalo-

sporins (such as cefadroxil, cephradine, cefuroxime and cephalothin). The mutant had, however, not achieved the maximum possible levels of resistance to third-generation cephalosporins and further attempts to re-model the active site of PBP 3 to give resistance to all of those cephalosporins that kill by inactivation of PBP 3 were carried out by selecting for increased resistance to cefoperazone, ceftazidime and cefotaxime. The fourth-level mutants obtained by selecting for resistance to the above compounds had only a slight further increase in their resistance to cephalosporins and had simultaneously regained the ability to form colonies at 43 °C. Attempts to produce fifth-level mutants by selecting for additional resistance to third-generation cephalosporins have been unsuccessful.

The properties of the series of mutants are shown in Table 1, while Table 2 shows the nucleotide substitutions that have occurred in the *pbpB* gene at each level of selection. Interestingly, the evolution of PBP3 in this series of mutants has occurred by amino-acid substitutions that alternately reduce the affinity of the enzyme for cephalosporins but de-stabilize the enzyme, and those that have little or no effect on affinity but which re-stabilize the enzyme.

The fourth-level mutant (SP1124) showed the maximum levels of resistance to almost all cephalosporins that can be achieved by reduction of the affinity of PBP3 (Table 1). The increase in resistance that has been gained in the latter mutant by four amino-acid substitutions in PBP3 would probably be sufficient to prevent successful therapy by most cephalosporins.

SP1124 showed cross-resistance to those cephalosporins that kill by inactivation of PBP3 but not to any of the penicillins or monobactams that were tested. As expected, the affinity of strain SP1124 PBP3 for benzylpenicillin was only slightly reduced from that of the parent strain (C600), whereas its affinity for cephalixin was greatly reduced (Fig. 1). It might be expected, *a priori*, that penicillins, monobactams and cephalosporins would have a substantial degree of overlap in their binding sites on the enzyme and it is surprising that the re-modelling of PBP3 to reduce its affinity for cephalosporins has had no apparent effect on the binding of the penicillins or monobactams. This suggests that there may, in fact, be little overlap between the binding sites for cephalosporins and the other classes of β -lactam antibiotics. Further evolution of the active site of PBP3 to exclude penicillins and monobactams as well as cephalosporins could probably be achieved, but we have not attempted to do this for safety reasons.

A plasmid that expressed the *pbpB* gene of the fourth-level mutant could transform a sensitive strain of *E. coli*, *Salmonella typhimurium* or *Klebsiella pneumoniae* to broad-spectrum cephalosporin resistance (Table 3). If resistant forms of PBP3 evolve in clinical isolates, the transfer of the altered PBP gene to a plasmid could therefore result in the infectious spread of resistance through *E. coli* and other species of the Enterobacteriaceae. The consequences of such events for the treatment of enterobacterial infections with β -lactam antibiotics would be considerable.

P.J.H. was in receipt of an MRC Biotechnology Studentship. We thank Dr J. K. Broome-Smith for helpful comments on the manuscript, Glaxo Group Research for oligonucleotides and Dr P. J. Cassidy of Merck Sharp and Dohme for ^3H -benzylpenicillin.

The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn

H. Jane Dyson, Keith J. Cross, Richard A. Houghten, Ian A. Wilson, Peter E. Wright & Richard A. Lerner

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037, USA

Many short synthetic peptides have now been shown to induce antibodies reactive with their cognate sequences in the intact folded protein¹⁻⁸. Aside from the usefulness of such antibodies as site-specific reagents, the frequency with which this recognition occurs has raised several theoretical issues, the central one being that of how an antibody to a short synthetic peptide, which represents one of the most disordered states of a site in a protein, can react with the more ordered version of the same sequence in the folded protein. This apparent paradox can be resolved if the target site on the protein approaches disorder or if the peptide in solution or on a carrier adopts, with significant frequency, a conformation compatible with that of the cognate site in the protein. Various studies already suggest that antigenic sites in proteins correspond to regions of high atomic mobility^{1,9-15}. We now show, using high-field nuclear magnetic resonance (NMR) spectroscopy, that a nonapeptide selected by several monoclonal antibodies as the immunodominant site of a 36-amino-acid immunogen (residues 75-110 of influenza virus haemagglutinin^{16,17}) adopts a highly populated type-II reverse-turn conformation in water. This suggests that in this case the antibodies have selected a sequence possessing a conformational preference. Apart from helping us to understand immunological recognition, anti-peptide antibodies may provide reagents of sufficient precision for an immunological approach to the problem of protein folding¹⁸⁻²³.

^1H -NMR spectroscopy provides a powerful method for determining molecular conformation in solution. In the present work, resonances were assigned to specific protons in the peptide using a combination of one- and two-dimensional methods. The conformation of the molecule was then examined by measuring nuclear Overhauser effects (NOEs), coupling constants and temperature and pH dependence of the NMR spectrum. The assignment procedure is illustrated in Fig. 1, which shows part of a two-dimensional scalar-correlated (COSY)²⁴⁻²⁶ spectrum. The cross-peaks establish unambiguously the coupling patterns (connectivities) for the resonances of each type of amino acid. Assignment to specific amino acids was based primarily on the observation of NOEs between resonances of sequentially adjacent amino acids. A detailed description of the assignments will be given elsewhere.

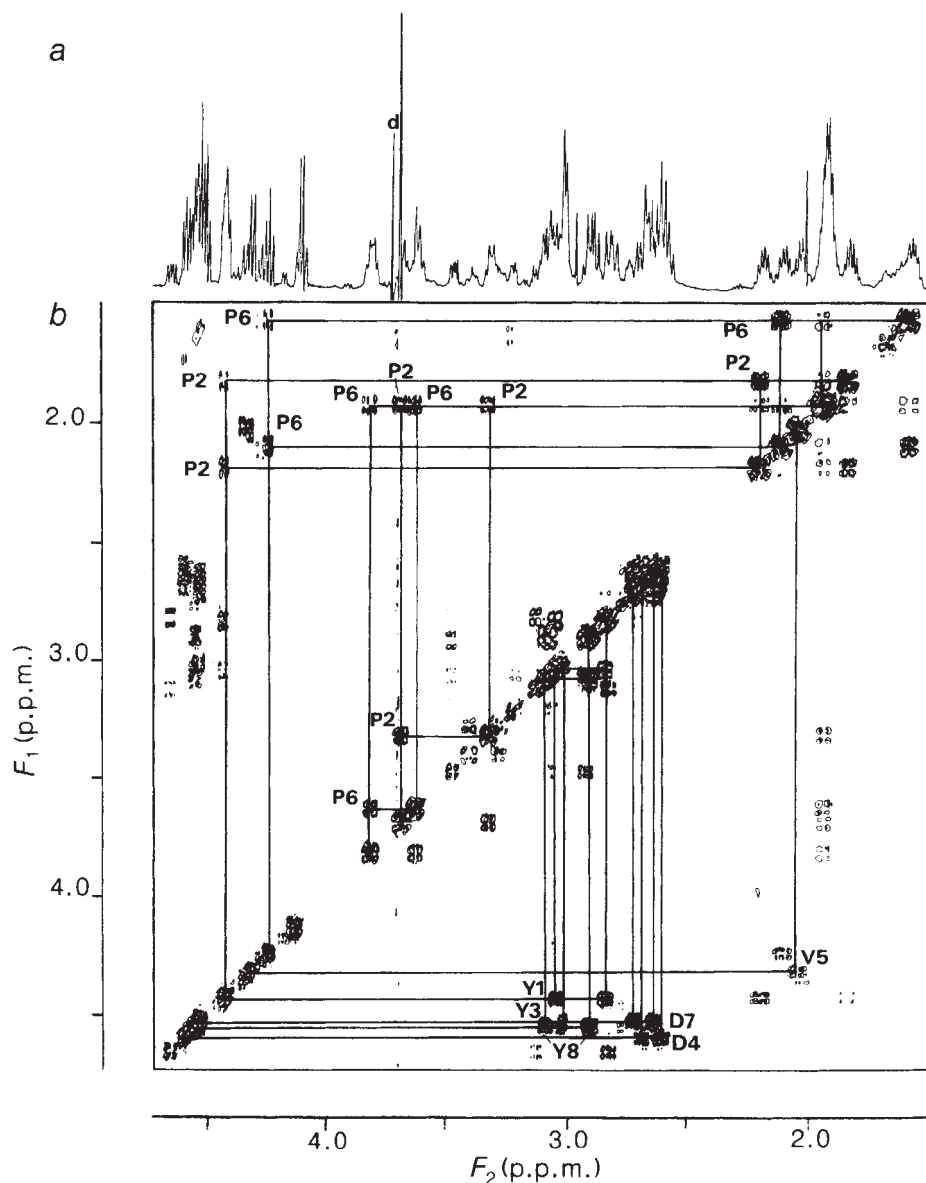
The NMR spectra of the peptide Tyr-Pro-Tyr-Asp-Val-Pro-Asp-Tyr-Ala are complicated by *cis-trans* isomerism²⁷ about the Tyr 1-Pro 2 peptide bond. This results in two distinct forms of the peptide which interconvert slowly on the NMR timescale. The relative intensities of peaks in the spectrum show that the two isomeric forms are present in the approximate *trans:cis* ratio of 2:1. This ratio appears to be little influenced by pH changes. Only the *trans* form appears to have demonstrable structure in water solution and thus is the subject of most of the following discussion.

The resonances of the exchangeable amide protons are absent from spectra of the peptide in D_2O . In H_2O , 12 amide proton resonances are observed, 6 from the *trans* and 6 from the *cis* form of the peptide (no NH proton resonances are expected for the N-terminal tyrosine or, of course, for the two prolines). The NH proton resonances were assigned from a COSY spectrum of the peptide in water solution (Fig. 2). Table 1 lists the assignments of the αCH and NH proton resonances. Unambiguous assignment of the aspartic acid spin systems is of

Received 19 August; accepted 11 October 1985.

1. Spratt, B. G. *J. gen. Microbiol.* **129**, 1247-1260 (1983).
2. Waxman, D. J. & Strominger, J. L. *A. Rev. Biochem.* **52**, 825-870 (1983).
3. Reynolds, P. E. *Br. med. Bull.* **40**, 3-10 (1984).
4. Parr, T. R. & Bryan, L. E. in *Antimicrobial Drug Resistance* (ed. Bryan, L. E.) 81-111 (Academic, New York, 1984).
5. Brown, F. J. & Reynolds, P. E. *FEBS Lett.* **122**, 275-278 (1980).
6. Hayes, M. V., Curtis, N. A. C., Wyke, A. & Ward, J. B. *FEMS Microbiol. Lett.* **10**, 119-122 (1981).
7. Hartman, B. J. & Tomasz, A. *J. Bact.* **158**, 513-516 (1984).
8. Rossi, L., Tonin, E., Cheng, Y. R. & Fontana, R. *Antimicrob. Ag. Chemother.* **27**, 828-831 (1985).
9. Zighelboim, S. & Tomasz, A. *Antimicrob. Ag. Chemother.* **17**, 434-442 (1980).
10. Parr, T. R. & Bryan, L. E. *Antimicrob. Ag. Chemother.* **25**, 747-753 (1984).
11. Dougherty, T. J., Koller, A. E. & Tomasz, A. *Antimicrob. Ag. Chemother.* **18**, 730-737 (1980).
12. Tipper, D. J. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **54**, 1133-1141 (1965).
13. Spratt, B. G. *Nature* **274**, 713-715 (1978).
14. Spratt, B. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2999-3003 (1975).
15. Sykes, R. B. *et al. Nature* **291**, 489-491 (1981).
16. Hedge, P. J. & Spratt, B. G. *Eur. J. Biochem.* **151**, 111-121 (1985).
17. Harder, H. J., Nikaïdo, H. & Matsushashi, M. *Antimicrob. Ag. Chemother.* **20**, 549-552 (1981).
18. Jaurin, B., Grundström, T. & Normark, S. *EMBO J.* **1**, 875-881 (1982).
19. Suzuki, H., Nishimura, Y. & Hirota, Y. *Proc. natn. Acad. Sci. U.S.A.* **75**, 664-668 (1978).
20. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
21. Coulondre, C. & Miller, J. H. *J. molec. Biol.* **117**, 525-567 (1977).
22. Hedge, P. J. & Spratt, B. G. *FEBS Lett.* **176**, 179-184 (1984).
23. Messing, J. & Vieira, J. *Gene* **19**, 269-276 (1982).
24. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
25. Nakamura, M., Maruyama, I. N., Soma, M., Kato, J.-I., Suzuki, H. & Hirota, Y. *Molec. gen. Genet.* **191**, 1-9 (1983).
26. Spratt, B. G. *Eur. J. Biochem.* **72**, 341-352 (1977).

Fig. 1 **a**, Portion of the 500-MHz ^1H -NMR spectrum of haemagglutinin peptide (YPYDVDPYA) at 278 K. All spectra were recorded using a Bruker AM500 spectrometer. Dioxan (d) was used as an internal standard, but the spectrum is referenced to TSS at 0 parts per 10^6 (p.p.m.). pH values in D_2O solution refer to uncorrected meter readings. Peptide concentration ~ 10 mM in D_2O , adjusted to pH 4.15 with 0.1 M NaOD and 0.1 M DCl solutions in D_2O . The residual HOD peak was suppressed by gated irradiation. Resolution was enhanced by lorentzian-to-gaussian weighting. **b**, Portion of the aliphatic region of a phase-sensitive, double quantum filtered COSY spectrum²⁶ of peptide. The spectrum was obtained at 278 K using the same solution as that used to obtain the spectrum in **a**. The residual HOD resonance was suppressed by gated irradiation for 2 s before acquisition. A total of 512 points were acquired in the t_1 domain, with 16×2 K free induction decays acquired for each t_1 value. The spectrum was Fourier transformed in both dimensions using a phase-shifted sine bell window function (shifted $\pi/4$ in t_1 and $\pi/8$ in t_2). Connectivities for the proline residues (P2 and P6) and the aspartate residues (D4 and D7), and the α - β connectivities for valine (V5) and tyrosine (Y1, Y3 and Y8) are indicated for the *trans* form only. All other cross-peaks arise from resonances of the *cis* form. **Methods.** The peptide was synthesized on a Biosearch Sam Two peptide synthesizer and purified by HPLC as reported previously³⁵. The sequence of the peptide corresponds to residues 98–106 of the HA1 chain of the H3 subtype X:47 (A/Victoria/3/75) haemagglutinin¹⁶. The immunological properties of this peptide have been studied in detail¹⁶. The affinity of the monoclonal anti-peptide antibodies for the peptide is in the range 1.6×10^5 – 3.5×10^6 M^{-1} and for the haemagglutinin monomer 8×10^5 – 4×10^6 M^{-1} (K. Bergmann, T. Klinger and I.A.W., unpublished observations).



particular importance. Assignments were therefore made by isotopic substitution; a peptide was synthesized in which the amide N of Asp 4 was substituted with $>90\%$ ^{15}N . This nucleus has spin $1/2$ and causes splitting of the resonance of the attached proton. The spectrum of the ^{15}N -substituted peptide (Fig. 3b) allowed unequivocal assignment of the Asp 4 and Asp 7 proton resonances (both *trans* and *cis* forms).

Several NMR parameters provide information on molecular conformation. The temperature dependence of amide proton resonances can be used to identify backbone hydrogen bonding. Except for Asp 4 (*trans* form), the NH proton resonances show the characteristic temperature dependence expected for random-coil peptides, where amide protons are completely exposed to solvent (Table 2). The temperature coefficients of amide proton resonances are expected to be -6 to -10 parts per 10^9 (p.p.b.) per K for random-coil peptides in H_2O (ref. 28). A value of -3.5 p.p.b. per K for Asp 4 NH indicates that this proton is protected from solvent to a considerable extent; it is also protected from solvent at low pH. Since its temperature coefficients are the same in all conditions studied, the *cis* form may be used in this system as an unfolded control. In the presence of 6 M urea, the temperature coefficient for the Asp 4 (*trans*) NH is close to that for the *cis* form (Table 2). We conclude that the *trans* form is unfolded in the presence of urea. In a small monomeric peptide, protection from solvent is probably due to

Table 1 Assignment of αCH and NH resonances of haemagglutinin peptide

		298 K		278 K	
		<i>trans</i>	<i>cis</i> *	<i>trans</i>	<i>cis</i> *
Tyr 1	αCH	4.44	3.506	4.442	3.493
Pro 2	αCH	4.44		4.441	
Tyr 3	αCH	4.575	4.669	4.534	4.668
	NH	7.788	8.191	7.992	8.365
Asp 4	αCH	4.624	4.60	4.607	
	NH	8.242	8.490	8.320	8.605
Val 5	αCH	4.365	4.383	4.333	4.361
	NH	7.937	7.896	8.117	8.188
Pro 6	αCH	4.284	4.313	4.259	4.291
Asp 7	αCH	4.57	4.57	4.523	
	NH	8.280	8.330	8.449	8.504
Tyr 8	αCH	4.55	4.57	4.560	
	NH	7.941	7.972	8.135	8.155
Ala 9	αCH	4.146	4.15	4.127	
	NH	7.898	7.928	8.038	8.052

Spectra obtained for 10 mM solutions of haemagglutinin peptide in D_2O (αCH resonances) and in 90% $\text{H}_2\text{O}/10\%$ D_2O (NH resonances). The pH of the solution in each case (uncorrected meter readings) was 4.10 ± 0.5 . Dioxan was used as internal standard.

* Due to spectral overlap and low intensity, not all of the resonances from the *cis* form of the peptide could be assigned.

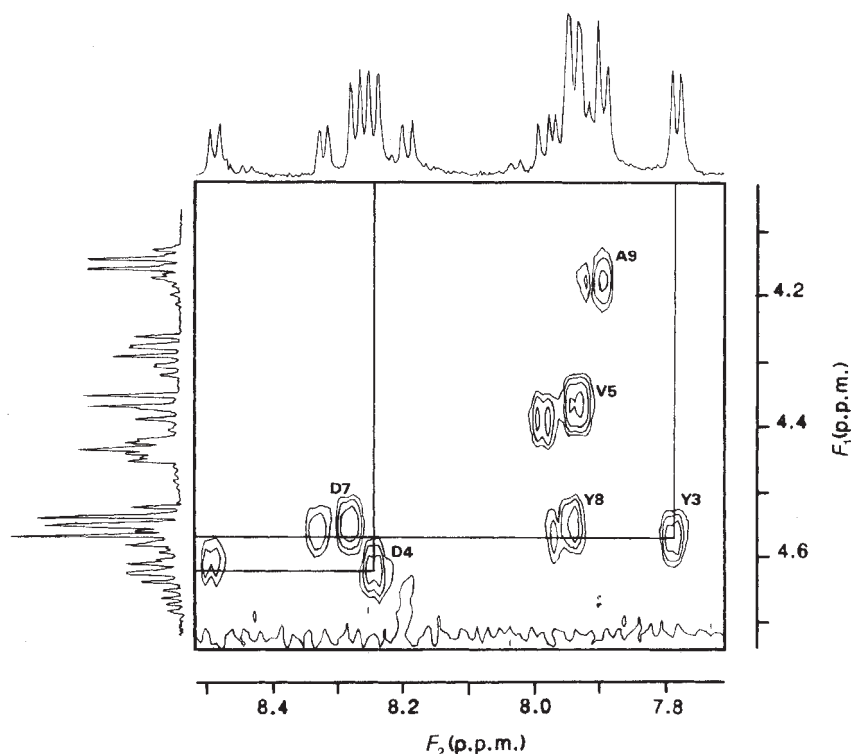


Fig. 2 Portion of the COSY spectrum of haemagglutinin peptide in H₂O solution. The spectrum was obtained using an 8 mM solution of peptide in 5% D₂O/95% H₂O at pH 4.15 and 296 K. A two-dimensional modification of the 'Jump-Return' sequence was used³⁶. 331 time domain points were acquired in t_1 , with 96 free induction decays acquired at each t_1 value. The spectrum was Fourier transformed in both dimensions using a sine bell window function. Cross-peaks between α CH and amide proton resonances are indicated for the *trans* form of the peptide only.

intramolecular hydrogen-bond formation. Reduced temperature coefficients for amide proton resonances have been used frequently to infer the presence of β -turns in peptides^{28,29}. These structures are usually observed in non-aqueous solvents such as dimethyl sulphoxide. In the present case, the temperature dependence of the Asp 4 (*trans*) amide proton resonance suggests participation in an intramolecular hydrogen bond in water solution. We note that the chemical shifts of the amide proton resonances are independent of peptide concentration over the range 1–10 mM, so that interactions between peptide molecules can be effectively discounted.

The existence of a β -bend in the haemagglutinin nonapeptide in water solution is confirmed by measurements of the NOEs. In a reverse turn formed by residues Tyr 1, Pro 2, Tyr 3 and Asp 4, there would be a hydrogen bond between the carbonyl group of Tyr 1 and the Asp 4 NH. Type-I, -II and -III β -turns are possible. In each case a small NOE is expected between the amide proton resonances of residues ($i+2$)(Tyr 3) and ($i+3$)(Asp 4)³⁰. A 2% NOE on the Tyr 3 NH is observed when the Asp 4 NH is irradiated. An NOE of similar size is seen on Asp 4 NH when Tyr 3 NH is irradiated (Fig. 3c). Type-II turns can be distinguished from type-I and -III turns by the presence of a strong NOE between the resonance of the α CH proton of residue ($i+1$) and the resonance of the amide proton of residue ($i+2$)³⁰. In the present case, a small NOE was observed on the α CH proton resonance of Pro 2 on irradiation of the Tyr 3 NH resonance. This NOE could not be readily quantitated because of the proximity of the Pro-2 α CH resonance to the large H₂O peak. In the reverse experiment, in which the α CH resonance of the proline was irradiated, an unequivocal NOE (9%) on the Tyr 3 NH resonance was observed (Fig. 3d). We conclude that, in water solutions of the peptide, a substantial population of conformers contains a type-II turn between residues 1 and 4. The β -turn is found only in the *trans* isomer; there is no evidence for any reverse-turn formation in the minor *cis* isomeric form.

The statistical weights of the three major side-chain rotamers³¹ of each amino-acid residue in the *trans* form of the haemagglutinin nonapeptide were calculated using values of the $^3J_{\alpha\text{H}-\beta\text{H}}$ coupling constants obtained from computer simulations of the strongly coupled $\text{C}\alpha\text{H}-\text{C}\beta\text{H}_2$ fragments of the spin systems. The rotamer populations for Tyr 1 are of particular note: it

appears that the *gauche* rotamer g^+g^- is not significantly populated in the *trans* form of the peptide. The absence of the g^+g^- rotamer is entirely consistent with the presence of a high proportion of molecules containing the β -turn structure, since this conformation would force energetically unfavourable steric crowding between the atoms involved in the turn and the aromatic ring of Tyr 1. We therefore conclude that the β -turn occurs in a high proportion of the conformational states of the *trans* form of the peptide in water solution.

The stable reverse turn reported here is in a peptide which is the site selected by the immune system as the immunodominant region of a larger peptide antigen. This observation bears significantly on several fundamental issues, including the problem of protein folding. Stable local structures, which may be different from those in the final folded protein, can be recognized by antibodies, and have been postulated as initiation sites in folding^{18,19,22,23}. Also, on theoretical grounds, reverse turns have been suggested as intermediates in protein folding²⁰. The experimental observation described here may thus link the immunological and theoretical approaches to protein folding.

Table 2 Temperature coefficients (p.p.b. per K) for NH resonances of haemagglutinin peptide

	pH 4.18		pH 1.50		pH 5.88 +6 M urea	
	<i>trans</i>	<i>cis</i>	<i>trans</i>	<i>cis</i>	<i>trans</i>	<i>cis</i>
Tyr	—	—	—	—	—	—
Pro	—	—	—	—	—	—
Tyr	-9.7	-8.0	-9.1	-7.3	-10.2	-7.6
Asp	-3.5	-5.7	-4.3	-5.8	-5.6	-6.3
Val	-8.3	-9.3	-7.3	-9.0	-9.0	-9.3
Pro	—	—	—	—	—	—
Asp	-8.3	-8.6	-8.1	-8.4	-8.7	-8.9
Tyr	-9.0	-9.1	-9.8	-9.0	-8.3	-8.3
Ala	-6.7	-6.7	-6.8	-6.8	-8.0	-7.9

Temperature coefficients were calculated by a linear least-squares fit to up to 12 data points for each resonance. The temperature coefficients recorded at all pH values above 4 were very similar, as expected since there are no deprotonations in the molecule between pH 4 and neutral pH. This suggests that the structure probably persists at neutral pH.

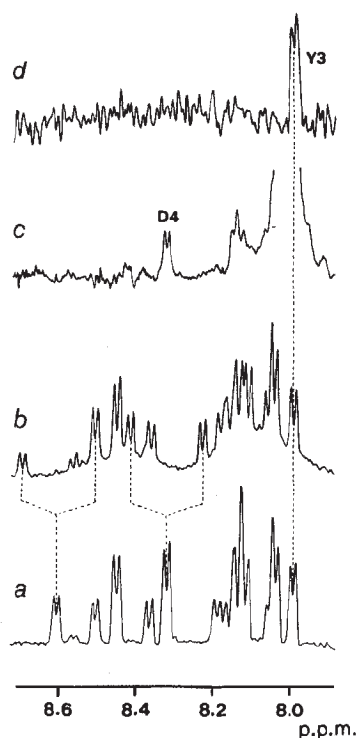


Fig. 3 *a*, Amide proton region of ^1H -NMR spectrum of the haemagglutinin peptide (10 mM in 90% H_2O /10% D_2O , adjusted to pH 4.10 with 0.1 M NaOH and 0.1 M HCl). The spectrum was obtained at 278 K using the Jump-Return sequence³⁷. Resolution was enhanced by lorentzian-gaussian transformation. 2,400 transients were collected. *b*, Amide proton region of spectrum of the haemagglutinin peptide with Asp 4 >90% enriched with ^{15}N . The spectrum shows splitting of the Asp 4 (*cis*) and Asp 4 (*trans*) NH resonances due to $^1J_{\text{NH}}$ coupling. The ^{15}N - β -benzyl-L-aspartate and ^{15}N -*tert*-butoxycarbonyl- β -benzyl aspartate intermediates for the synthesis of this peptide were prepared from >90% L-aspartic- ^{15}N acid (MSD Isotopes) by standard procedures^{38,39}. Peptide concentration was ~2 mM in 90% H_2O /10% D_2O at pH 4.07 at 278 K. *c*, NOE difference spectrum with 1.5-s gated irradiation of Tyr 3 NH resonance at 7.99 p.p.m. Total numbers of transients accumulated, 1,234. The free induction decay was subjected to exponential multiplication before Fourier transformation. *d*, NOE difference spectrum with 1.5-s gated irradiation of Pro-2 αCH resonance at 4.44 p.p.m. Total number of transients accumulated, 952. Resolution was enhanced by lorentzian-gaussian transformation. For the NOE experiments, spectra with the decoupler on- and off-resonance were interleaved (8 transients of each) and stored on the disk.

The main immunological question raised by this work is whether the presence of a stable reverse turn in a synthetic immunogen in water is important for recognition of the cognate sequence in the folded protein. A related question involves the extent to which antibody binding influences the conformation of the protein and the peptide. Previous studies using anti-protein antibodies have suggested that interaction with the antibody may itself introduce order into a peptide^{22,23,32-34}. In the present situation, the conformation of the peptide in solution and the cognate site in the crystal structure of the protein are different, suggesting that one or both must be altered in the antibody-binding site. Since we are dealing with a monoclonal antibody, the most direct approach to answering such questions is to determine the conformation of the synthetic immunogen when bound to the antibody. These studies are in progress, using both crystallographic and NMR methodologies.

Pending results from these studies, we can state our current view of how the immunological issues can be resolved. It seems that there may be a complementarity between the view that the site in the protein to which antibody binds is mobile and the postulate that the immunogenic peptide may be partially

ordered. In keeping with the diversity of the immune system, we would not expect all anti-peptide antibodies to be the same. Sometimes the balance will be towards disorder in the protein whereas at other times more order in the immunogen will be required.

We thank Dr Mark Rance for help with the NMR experiments, John Ostresh for synthesis of the ^{15}N -peptide and Gail Donnan for technical assistance. We acknowledge financial support from NIH grant AI19499 (to R.A.L. and I.A.W.). This is publication MB-3969 from Scripps Clinic.

Note added in proof: NMR studies of a peptide corresponding to an antigenic region of myelin basic protein have now been reported⁴⁰.

Received 19 June; accepted 3 October 1985.

1. Lerner, R. A. *Adv. Immun.* **36**, 1-44 (1984).
2. Lerner, R. A. *Nature* **299**, 592-596 (1982).
3. Enea, V. *et al. Science* **225**, 628-630 (1984).
4. Ballou, W. R. *et al. Science* **228**, 996-999 (1985).
5. Chow, M., Yabrov, R., Bittle, J., Hogle, J. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 910-914 (1985).
6. Kris, R. M. *et al. Cell* **40**, 619-625 (1985).
7. Lamb, R. A., Zebede, S. L. & Richardson, C. D. *Cell* **40**, 627-633 (1985).
8. Schmidt, M. A., O'Hanley, P. & Schoolnik, G. K. *J. exp. Med.* **161**, 705-717 (1984).
9. Artymuk, P. J. *et al. Nature* **280**, 563-568 (1979).
10. Moore, G. R. & Williams, R. J. P. *Eur. J. Biochem.* **103**, 543-550 (1980).
11. Westhof, E. *et al. Nature* **311**, 123-126 (1984).
12. Tainer, J. A. *et al. Nature* **312**, 127-133 (1984).
13. Williams, R. J. P. & Moore, G. R. *Trends biochem. Sci.* **10**, 96-97 (1985).
14. Tainer, J. A., Getzoff, E. D., Paterson, Y., Olson, A. J. & Lerner, R. A. *Rev. Immun.* **3**, 501-535 (1985).
15. Hirayama, A., Takagaki, Y. & Karush, F. *J. Immun.* **134**, 3241-3247 (1985).
16. Wilson, I. A. *et al. Cell* **37**, 767-778 (1984).
17. Wilson, I. A., Wiley, D. C. & Skehel, J. J. *Nature* **289**, 366-373 (1981).
18. Sachs, D. H., Schechter, A. N., Eastlake, A. & Anfinsen, C. B. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3790-3794 (1972).
19. Sachs, D. H., Schechter, A. N., Eastlake, A. & Anfinsen, C. B. *J. Immun.* **109**, 1300-1310 (1972).
20. Anfinsen, C. B. & Scheraga, H. A. *Adv. Protein Chem.* **29**, 205-300 (1975).
21. Wetlauffer, D. B. *Adv. Protein Chem.* **34**, 61-92 (1981).
22. Teale, J. M. & Benjamin, D. C. *J. Biol. Chem.* **251**, 4609-4615 (1976).
23. Celada, F., Fowler, A. V. & Zabin, I. *Biochemistry* **17**, 5156-5160 (1978).
24. Aue, W. P., Bartholdi, E. & Ernst, R. R. *J. chem. Phys.* **64**, 2229-2246 (1976).
25. Bax, A. & Freeman, R. *J. magn. Reson.* **44**, 542-561 (1981).
26. Rance, M. *et al. Biochem. biophys. Res. Commun.* **117**, 479-485 (1983).
27. Grathwohl, C. & Wuthrich, K. *Biopolymers* **20**, 2623-2633 (1981).
28. Deslauriers, R. & Smith, I. C. P. in *Biological Magnetic Resonance* Vol. 2 (eds Berliner, L. J. & Reuben, J.) 243-344 (Plenum, New York, 1980).
29. Urry, D. W. & Ohnishi, M. in *Spectroscopic Approaches to Biomolecular Conformation* (ed. Urry, D. W.) 263-300 (American Medical Association, Chicago, 1970).
30. Shenderovich, M. D., Nikiforovich, G. V. & Chipens, G. I. *J. magn. Reson.* **59**, 1-12 (1984).
31. Bystron, V. F. *Prog. NMR Spectrosc.* **10**, 41-81 (1976).
32. Crumpton, M. J. & Small, P. A. *J. molec. Biol.* **26**, 143-146 (1967).
33. Conway-Jacobs, A., Schechter, B. & Sela, M. *Biochemistry* **9**, 4870-4875 (1970).
34. Schechter, B., Conway-Jacobs, A. & Sela, M. *Eur. J. Biochem.* **20**, 321-324 (1971).
35. Houghten, R. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5131-5135 (1985).
36. Guittet, E., Delsuc, M. A. & Lallemand, J. Y. *J. Am. chem. Soc.* **106**, 4278-4279 (1984).
37. Plateau, P. & Gueron, M. *J. Am. chem. Soc.* **104**, 7311-7312 (1982).
38. Benoiton, L. *Can. J. Chem.* **40**, 570-572 (1962).
39. Itoh, M., Hagiwara, D. & Kamiya, T. *Tetrahedron Lett.* **49**, 4393-4394 (1975).
40. Mendz, G. L. & Moore, W. J. *Biochem. J.* **229**, 305-313 (1985).

Identification of kinesin in sea urchin eggs, and evidence for its localization in the mitotic spindle

J. M. Scholey, M. E. Porter, P. M. Grissom & J. R. McIntosh

Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Campus Box 347, Boulder, Colorado 80309, USA

To understand the molecular basis of microtubule-associated motility during mitosis^{1,2}, the mechanochemical factors that generate the relevant motile force must be identified. Myosin, the ATPase that interacts with actin to produce the force for muscle contraction and other forms of cell motility³, is believed to be involved in cytokinesis but not in mitosis⁴⁻⁷. Dynein, the mechanochemical enzyme that drives microtubule sliding in eukaryotic cilia and flagella^{8,9}, has been identified in the cytoplasm of sea urchin eggs¹⁰⁻¹⁹, but the evidence that it is involved in

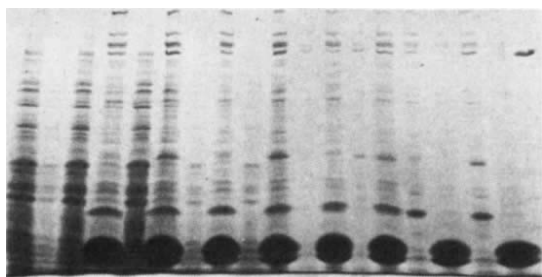


Fig. 1 Specific AMPPNP-dependent co-pelleting of the 134K polypeptide with unfertilized egg cytoplasmic microtubules. The figure shows gels of the supernatants (lanes 1, 3, 5) and pellets (lanes 2, 4, 6) obtained from different experiments. Lanes 1, 2, plus AMPPNP plus colchicine (to block microtubule assembly); lanes 3, 4, no additions; lanes 5, 6, plus AMPPNP. Note that the 134K protein is only recovered in pellet 6, containing AMPPNP but not colchicine. Control microtubule (lane 4) and 'AMPPNP' microtubule (lane 6) pellets were washed to yield the supernatants and pellets shown in lanes 7, 8 and 9, 10, respectively. These pellets were subsequently extracted with 5 mM Mg-ATP, yielding supernatants and pellets shown in lanes 11 and 12, and 13 and 14, respectively. Note that the putative cytoplasmic dynein heavy chain (Dyn.) and the 134K polypeptide remain bound to the AMPPNP-microtubules through the washing step, but are both partially extracted by ATP (lane 13). The rest of the MAPs were extracted from the control and AMPPNP microtubules by differential centrifugation in 0.5 M NaCl plus 0.1 mM ATP; the corresponding supernatant and pellets are shown in lanes 15 and 16 and 17 and 18, respectively. Lane 19 is sea urchin sperm flagella 21S dynein. Tub., tubulin; 80K, an 80,000-M_r MAP.

Methods. Unfertilized sea urchin eggs were washed, dejellied and homogenized in 'extraction buffer' (0.1 M PIPES, 2.5 mM Mg(CH₃COO)₂, 5 mM EGTA, 0.1 mM EDTA, 0.9 M glycerol, 0.1 mM phenylmethylsulphonyl fluoride, 1 µg ml⁻¹ pepstatin, 1 µg ml⁻¹ leupeptin, 10 µg ml⁻¹ aprotinin, 0.5 mM dithiothreitol, pH 6.9), then centrifuged to yield the cytoplasmic extract¹³ in which microtubules were assembled by incubation with taxol and GTP^{13,28} for 20 min. 0.1 M AMPPNP in extraction buffer was then added to a final concentration of 5 mM to one aliquot for a further 10 min, while a control aliquot was incubated with an equal volume of extraction buffer. One sample of extract was treated with 100 µM colchicine instead of taxol to inhibit microtubule assembly, then supplemented with 5 mM AMPPNP. These extracts were centrifuged through a 15% sucrose cushion¹³. The SDS gel system used here contained 6% acrylamide plus 0.06% bis-acrylamide. The low concentration of cross-linker allows easy visualization of the dynein heavy chain. Other biochemical procedures were performed as described previously¹³.

cytoplasmic microtubule-based motility (rather than serving as a precursor for embryonic cilia) is equivocal. Microtubule-associated ATPases have been prepared from other tissues (reviewed in ref. 12), but their role in cytoplasmic motility is also unknown. Recent work on axoplasmic transport, however, has led to the identification of a novel mechanochemical protein called kinesin²⁰, which is thought to generate the force for moving vesicles along axonal microtubules²⁰⁻²⁷. These results suggest that kinesin may also be a mechanochemical factor for non-axoplasmic forms of microtubule-based motility, such as mitosis. We describe here the identification and isolation of a kinesin-like protein from the cytoplasm of sea urchin eggs. We present evidence that this protein is localized in the mitotic spindle, and propose that it may be a mechanochemical factor for some form of motility associated with the mitotic spindle.

Extracts of squid axoplasm will support a microtubule-based, ATP-dependent movement of vesicles²⁰⁻²⁷. This motility is driven by a translocator that exists primarily in soluble form, with a fraction serving to cross-link microtubules to the moving vesicles²⁴. The non-hydrolysable ATP analogue, adenylyl imidodiphosphate (AMPPNP) 'freezes' vesicle motion on microtubules^{22,24,26}, suggesting that the analogue induces strong bind-

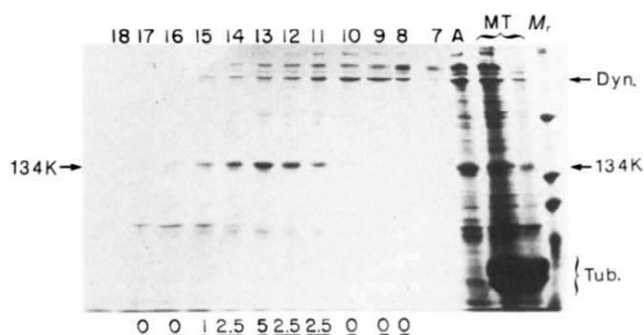


Fig. 2 Biogel A5M chromatography of sea urchin egg MAPs. Microtubules (MT, two different loadings) prepared in AMPPNP were washed in extraction buffer supplemented with 75 mM NaCl. The 134K polypeptide (134K) and putative dynein heavy chain (Dyn.) were extracted from the microtubules by differential centrifugation in buffer containing 0.1 M NaCl/10 mM MgSO₄/7.5 mM ATP (lane A), then fractionated on a 1.2 × 20-cm column of Biogel A5M. Fractions of 50 drops were collected, then 40-µl aliquots were analysed on SDS gels (lanes 7-18). Fraction 8 represents the void volume (V₀) and fraction 17 is the included volume (V_i). Aliquots (100 µl) of each fraction were frozen in liquid nitrogen, then transported on dry ice by Federal Express to the Marine Biological Laboratories, Woods Hole, where R. D. Vale and M. P. Sheetz analysed them for their ability to induce microtubules to move relative to other microtubules 'in vitro' as described by Vale *et al.*²⁰ for squid and bovine brain kinesin. The numbers beneath each column fraction represent the maximum dilution factor at which motility-inducing activity can still be detected. 0 = No motility observed. Vale and Sheetz did not know the polypeptide composition of the samples before assay. Note that motility-inducing activity co-elutes with the 134K polypeptide, the most active fraction being 13 (K_{av} ≈ 0.56).

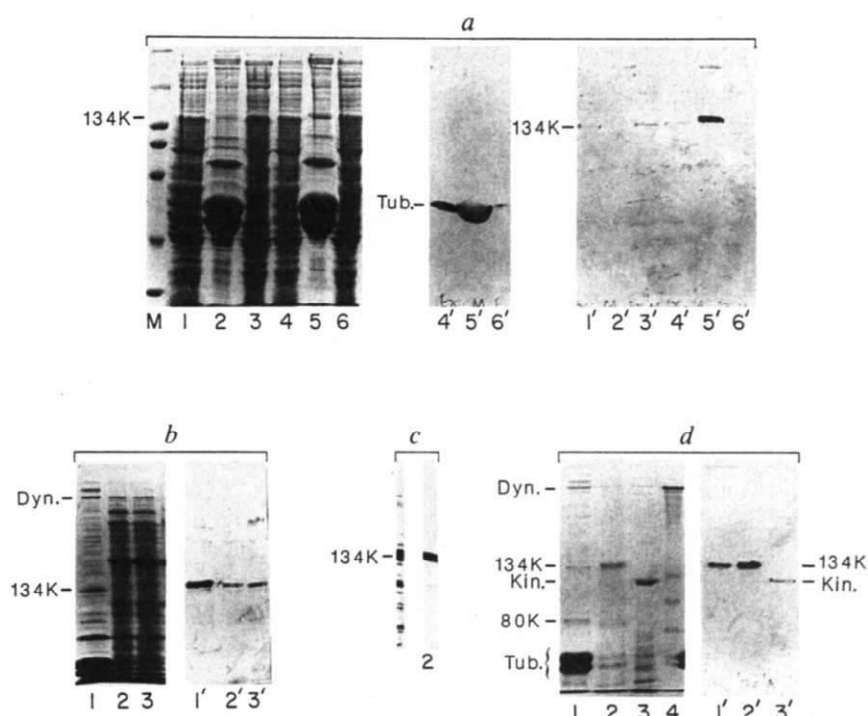
ing between the translocators and microtubules, thereby driving the translocators from a soluble to a bound form. We reasoned that AMPPNP might enhance microtubule binding of the translocators present in cytoplasmic extracts, so that microtubules prepared in AMPPNP might serve as an affinity matrix for identifying and isolating these mechanochemical proteins.

Taxol-assembled microtubules were therefore prepared from cytoplasmic extracts of unfertilized sea urchin eggs^{13,28} in the presence and absence of 1-5 mM AMPPNP. The amount of various microtubule-associated proteins (MAPs) co-sedimenting with tubulin was determined by densitometry of SDS-polyacrylamide gels (Fig. 1). We previously reported that a cytoplasmic dynein-like protein present in sea urchin egg extracts exists in both soluble and microtubule-associated forms^{13,14}. Only a small increase in the amount of the putative dynein heavy chain was observed in microtubules pelleted in the presence compared with the absence of AMPPNP (0.6 compared with 0.4 mol per 100 mol tubulin). Furthermore, we observed no significant increase in the ATPase activity which co-pelleted with the microtubules in the presence of AMPPNP (microtubules + AMPPNP: specific activity ~10 nmol per min per mg, total activity 31 nmol per min; microtubules - AMPPNP: specific activity ~13 nmol per min per mg, total activity 36 nmol per min). These results suggest that cytoplasmic dynein interacts with microtubules differently from the squid axoplasmic vesicle translocator^{22,24,26}; they do not support the hypothesis that cytoplasmic dynein is the motor for vesicle transport.

Addition of AMPPNP did, however, greatly enhance the amount of a polypeptide of relative molecular mass (M_r) 134,000 (134K) which co-sedimented with microtubules (Fig. 1) (~8 mol 134K per 100 mol tubulin compared with ~1 mol 134K per 100 mol tubulin without AMPPNP). Identical results were obtained when an egg extract was treated with 1 mM ADP plus 100 µM Na₃VO₄, or when the ATP concentration in the extract was depleted by incubation with apyrase. The 134K polypeptide

Fig. 3 Antibody to the sea urchin egg 134K polypeptide stains a 134K polypeptide in immunoblots of cytoplasmic microtubules prepared in AMPPNP (*a*) and isolated mitotic spindles (*b*). It also reacts with the 110K polypeptide of squid kinesin (*d*). *a*, 7.5% SDS gels of microtubules prepared in the presence and absence of AMPPNP (Fig. 1). Lanes 1–6, Coomassie blue-stained gel; lane M shows M_r markers: 200,000; 116,000; 96,000; 68,000; 44,000; and 29,000. Lanes 1, 4, cytoplasmic extracts; lane 2, microtubules prepared in the absence of AMPPNP; lane 5, microtubules prepared in the presence of AMPPNP; lanes 3, 6, the corresponding supernatants. Immunoblots of corresponding lanes (1'–6') probed with anti-tubulin¹³ (Tub.) and 134K antibody are also shown. The 134K antibody reacts with a 134K polypeptide in microtubules pelleted in the presence (lane 5') but not in the absence (lane 2') of AMPPNP. *b*, The 6% polyacrylamide gels (lanes 1–3) and corresponding anti-134K immunoblots (lanes 1'–3') show microtubules prepared in the presence of AMPPNP (lanes 1, 1') and mitotic spindles which were isolated³² in the presence (lanes 2, 2') and absence (lanes 3, 3') of AMPPNP. An immunoreactive 134K polypeptide is present in spindles isolated either way. *c*, Blot-affinity purification^{29–31} of 134K antibodies. Lane 1, a nitrocellulose strip of 'AMPPNP' microtubules probed with whole 134K antiserum. Lane 2 was probed with affinity-purified antibodies eluted from 134K polypeptide/nitrocellulose strips²⁹. *d*, Cross-reaction of sea urchin egg 134K antibody with squid axoplasmic kinesin. The Coomassie blue-stained gel (lanes 1–4) and corresponding anti-134K immunoblot (lanes 1'–3') show microtubules prepared from fertilized sea urchin eggs in the presence of AMPPNP (lanes 1, 1'), sea urchin egg kinesin (lanes 2, 2') and purified squid axoplasmic kinesin (lanes 3, 3'). Lane 4, a 21S dynein marker. The 134K urchin polypeptide and the 110K squid kinesin polypeptide (Kin.) are indicated.

Methods. Two rabbits were injected at 2–3-week intervals with 134K polypeptide excised from stained–destained SDS gels of proteins extracted from 'AMPPNP-microtubules' in ATP + NaCl, over a period of 4 months. No 134K antibodies were detected in preimmune sera from either rabbit, but 134K antibodies could be detected on immunoblots at dilution 1:1,000 after the first and all subsequent booster injections.



did not pellet from AMPPNP-treated cytoplasmic extracts of unfertilized sea urchin eggs in which microtubule assembly was inhibited by treatment with colchicine (Fig. 1). Therefore, in the presence of AMPPNP, the recovery of this polypeptide in the microtubule pellet depends on its binding to the microtubules. The polypeptide composition of taxol-assembled microtubules prepared in the presence of AMPPNP was unchanged by the addition of 0.1% Triton X-100, suggesting that the 134K polypeptide is not associated with membranous vesicles. Similar results were obtained when cytoplasmic extracts of fertilized eggs were used to prepare microtubules.

Like the cytoplasmic dynein heavy chain, the 134K polypeptide was not released from the microtubules by differential centrifugation in the original egg extraction buffer containing 75 mM NaCl, but it was partially extracted using 5–10 mM Mg-ATP, and completely extracted using 0.5 M NaCl or 0.1 M NaCl, 7.5–10 mM ATP plus 10 mM MgSO₄ (Figs 1, 2). As the latter procedure gave the most efficient and specific extraction of 134K protein, this material was used for further purification by gel filtration chromatography (Fig. 2).

The sea urchin egg 134K protein resembles kinesin²⁰ in several respects. (1) The binding of both proteins to microtubules is greatly enhanced by AMPPNP; (2) they are extracted from microtubules in 0.1 M NaCl plus 10 mM Mg-ATP; (3) their elution behaviour on gel filtration columns suggests that they contain multiple subunits; (4) fractions containing the sea urchin egg 134K polypeptide cause microtubules to move relative to other microtubules, and beads to translocate along microtubules 'in vitro' in a manner very similar to squid kinesin (Fig. 2). These results strongly suggest that the 134K polypeptide represents sea urchin egg kinesin.

Specific antisera were raised against the sea urchin 134K polypeptide by injecting rabbits with electrophoretically purified 134K polypeptide. These antibodies stained the 134K polypep-

tide present in immunoblots of microtubules pelleted from egg extracts in the presence of AMPPNP (Fig. 3). This 134K antigen was enriched in microtubule pellets relative to the original cytoplasmic extracts, but it was not observed in immunoblots of pellets prepared in the absence of AMPPNP (Fig. 3). To analyse the immunological relatedness of the sea urchin 134K polypeptide and squid axoplasmic kinesin, blots of both proteins were probed with the anti-134K antiserum (Fig. 3d) or with blot-affinity-purified 134K antibody (Fig. 3c)^{29–31}. The 134K antibody exhibited a clear cross-reaction with the major polypeptide component (M_r ~110K) of purified squid axoplasmic kinesin. Furthermore, antibodies affinity-purified from the 134K antisera on blots of the squid kinesin 110K polypeptide also cross-reacted with the sea urchin 134K polypeptide (data not shown). These data strongly suggest that the sea urchin egg 134K polypeptide represents kinesin. It is interesting that the M_r s of squid axoplasmic kinesin (110,000; ref. 20) and sea urchin egg kinesin (~134,000) are clearly different (Fig. 3).

The observation that a kinesin-like protein is associated with microtubules from the cytoplasm of sea urchin eggs, where chromosome movement is a major function of these structures, suggested that kinesin might be a component of the mitotic spindle. Immunoblots of mitotic spindles isolated from sea urchin embryos³² showed that antibodies to the 134K protein recognized a polypeptide of appropriate M_r in spindles isolated with or without AMPPNP. It is uncertain why AMPPNP is not required for association of the 134K protein with spindle isolates. However, since the isolation protocol involves egg lysis in ~100 volumes of buffer containing no ATP, the resulting depletion of ATP concentration may be sufficient to enhance microtubule-binding by the 134K protein.

The localization of the 134K polypeptide in fixed mitotic sea urchin eggs was investigated by immunoperoxidase light microscopy using blot-affinity-purified 134K antibodies (Fig. 4).

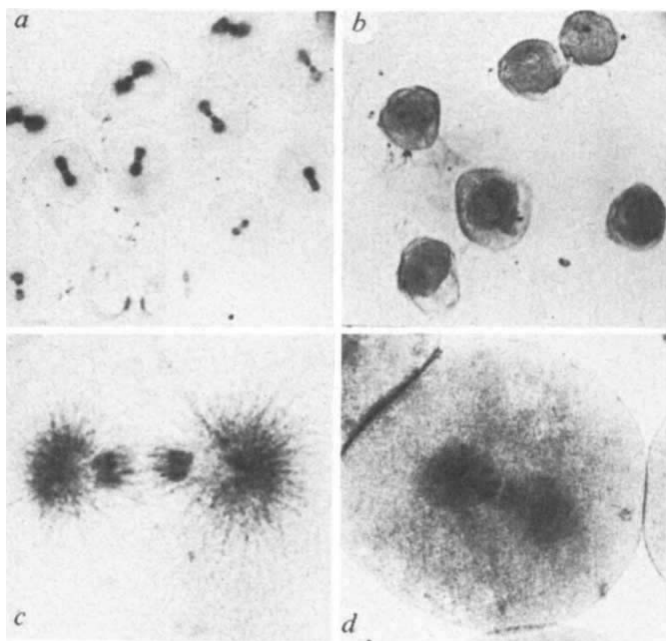


Fig. 4 Immunoperoxidase light microscopic localization of the 134K kinesin-like protein in dividing sea urchin embryos. *a*, A low-magnification view of lysed cells stained with blot affinity-purified 134K antibody, showing clear spindle staining. Whole immune sera (diluted 1:100–1:1,000) from both rabbits also exhibited clear spindle staining, whereas corresponding pre-immune sera exhibited no microtubule staining. *b*, Control lacking primary antibody. *c*, High-magnification view of a lysed cell stained with blot-affinity-purified 134K antibody, showing clear localization of the 134K polypeptide in the spindle fibres. *d*, Staining of an unlysed dividing blastomere. Note the concentration of the 134K protein in the mitotic spindle region. The approximate magnifications are: *a*, *b*, $\times 100$; *c*, $\times 700$; *d*, $\times 500$.

Methods. Sea urchin eggs were fertilized, allowed to progress to metaphase and fixed in 90% methanol plus 50 mM EGTA with or without prior lysis in 0.5% Nonidet P-40 (ref. 37). The fixed cells were incubated with blot-affinity-purified 134K antibodies and processed for peroxidase-antiperoxidase (PAP) staining and bright-field light microscopy.

The antibodies were checked for monospecificity on immunoblots of whole egg homogenates before use. Spindle staining was obvious (Fig. 4*a*). Observations of unlysed metaphase eggs stained with the affinity-purified 134K antibody revealed that the corresponding antigen is present throughout the cytoplasm, but is highly concentrated in the mitotic spindle (Fig. 4*d*). This pattern may reflect the existence of pools of soluble kinesin (accounting for background staining) and microtubule-associated kinesin (producing the spindle staining). Tubulin is believed to exhibit a similar distribution in dividing sea urchin eggs³³. When the blastomeres were lysed with detergent before fixation, clear staining of both chromosome-to-pole and astral fibres was evident (Fig. 4*c*). Several control experiments confirmed that staining was specific for 134K antibodies. (1) No microtubule staining was observed with preimmune sera or when the cells were incubated with secondary antibodies alone (Fig. 4*b*). (2) When 134K antibodies were preincubated with the 134K protein prepared by gel filtration chromatography (Fig. 2), no microtubule staining was observed by subsequent immunoperoxidase light microscopy. (3) Preabsorption of the 134K antibody with bovine brain tubulin did not affect staining. (4) Antibodies prepared from the 134K antisera by blot-affinity purification against the squid kinesin 110K polypeptide also stained the spindle fibres by immunoperoxidase light microscopy. These results strongly suggest that spindle staining is not due to a contaminating 134K polypeptide in our sea urchin microtubule preparations, but rather to the presence of a kinesin-like protein in the spindle fibres.

We propose that the 134K polypeptide represents a cytoplasmic form of kinesin which is associated with the mitotic spindle in dividing sea urchin embryos. The observations that kinesin induces microtubule-based motility *in vitro* suggest that it may also promote motility within the cell, so it is tempting to speculate that the sea urchin protein performs a mechanochemical function within the mitotic spindle. For example, it may cause microtubules to slide relative to other microtubules or to some other components of the spindle, or it may cause chromosomes to be translocated along microtubules. Alternatively, cytoplasmic kinesin may not be directly involved in chromosome movement, but may instead be involved in moving 'particles or states' along microtubules in the metaphase half-spindle (see, for example, ref. 34). To distinguish between such possibilities and to elucidate the function of kinesin within the cell, other techniques such as antibody microinjection into living cells³⁵ are being used.

A predicted property of kinesin is ATPase activity. Interestingly, since this manuscript was submitted, Brady reported³⁶ that the AMPPNP-dependent binding of a 130K polypeptide to chick brain microtubules accompanies an increase in the ATPase activity of the microtubules. However, the exact relationship between this ATPase and kinesin is not yet established.

We thank Drs Ron Vale and Mike Sheetz of the NIH NINCDS group at the Marine Biological Laboratories, Woods Hole, for providing squid axoplasmic kinesin and for performing motility assays on the sea urchin egg protein. Their results and the characterization of the sea urchin protein will be described elsewhere. We also thank numerous colleagues for their interest in this work, Larry Harwood for photography and Cathy Inouye and Karen Brown for secretarial assistance. J.M.S. particularly thanks Bonnie Neighbors for advice on immunoperoxidase light microscopy. Taxol was a gift of Dr M. Suffness (National Products Branch, NCI). J.M.S. was supported by an MRC Travelling Fellowship and a British-American exchange fellowship from the British Heart Foundation and American Heart Association. M.E.P. was supported by a fellowship from the Helen Hay Whitney Foundation. This work was supported by NIH grant GM 30213 and American Cancer Society grant BC-498 to J.R.M.

Received 14 August; accepted 9 October 1985.

- McIntosh, J. R. *Trends biochem. Sci.* **100**, 195–198 (1984).
- Schliwa, M. in *Cell and Muscle Motility* Vol. 5 (ed. Shay, J. W.) 1–82 (Plenum, New York, 1984).
- Harrington, W. F. & Rodgers, M. E. *Rev. Biochem.* **53**, 35–73 (1984).
- Fujiwara, K. & Pollard, T. D. *J. Cell Biol.* **71**, 848–875 (1976).
- Yumura, S. & Fukui, Y. *Nature* **314**, 194 (1985).
- Mabuchi, I. & Okuno, M. *J. Cell Biol.* **74**, 251–263 (1977).
- Kiehart, D. P., Mabuchi, I. & Inoue, S. *J. Cell Biol.* **94**, 165–178 (1982).
- Johnson, K. A. *Rev. biophys. Chem.* **14**, 161–188 (1985).
- Gibbons, I. R. *J. Cell Biol.* **91**, 1075–1245 (1981).
- Pratt, M. M. *Dev. Biol.* **74**, 364–378 (1980).
- Hisanaga, S. I. & Sakai, H. *J. Biochem., Tokyo* **93**, 87–98 (1983).
- Pratt, M. M. *Int. Rev. Cytol.* **87**, 83–105 (1984).
- Scholey, J. M., Neighbors, B., McIntosh, J. R. & Salmon, E. D. *J. biol. Chem.* **259**, 6516–6525 (1984).
- Dinenberg, A. S., McIntosh, J. R. & Scholey, J. M. *Ann. N.Y. Acad. Sci.* (in the press).
- Hollenbeck, P. J., Supryniewicz, F. & Cande, W. Z. *J. Cell Biol.* **99**, 1251–1258 (1984).
- Penningroth, S. M. *et al. Cell Motility* **5**, 61–75 (1985).
- Asai, D. & Wilson, L. *J. biol. Chem.* **260**, 699–702 (1985).
- Pratt, M., Otter, T. & Salmon, E. D. *J. Cell Biol.* **86**, 738–745 (1980).
- Piperno, G. *J. Cell Biol.* **98**, 1842–1850 (1984).
- Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **42**, 39–50 (1985).
- Allen, R. D. *et al. J. Cell Biol.* **100**, 1736–1752 (1985).
- Brady, S., Allen, R. D. & Lasek, R. *Cell Motility* **5**, 81 (1985).
- Vale, R. D., Schnapp, B. J., Reese, T. S. & Sheetz, M. P. *Cell* **40**, 449–454 (1985).
- Vale, R. D., Schnapp, B. J., Reese, T. S. & Sheetz, M. P. *Cell* **40**, 553–569 (1985).
- Schnapp, B. J., Vale, R. D., Sheetz, M. P. & Reese, T. S. *Cell* **40**, 455–462 (1985).
- Lasek, R. J. & Brady, S. T. *Nature* **316**, 645–647 (1985).
- Gilbert, S. P., Allen, R. D. & Sloboda, R. D. *Nature* **315**, 245–248 (1985).
- Vallee, R. B. & Bloom, G. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6259–6263 (1983).
- Olmsted, J. B. *J. biol. Chem.* **256**, 11955–11957 (1981).
- Talian, J. C., Olmstead, J. B. & Goldman, R. D. *J. Cell Biol.* **97**, 1277–1282 (1983).
- Smith, D. E. & Fisher, P. A. *J. Cell Biol.* **99**, 20–23 (1984).
- Salmon, E. D. *Meth. Cell Biol.* **25**, 69–105 (1982).
- Salmon, E. D., Leslie, R. J., Saxton, W. M., Karow, M. L. & McIntosh, J. R. *J. Cell Biol.* **99**, 2165–2174 (1984).
- Allen, R. D., Bajer, A. & LaFountain, J. *J. Cell Biol.* **43**, 4a (1969).
- Izant, J. G., Weatherbee, J. A. & McIntosh, J. R. *J. Cell Biol.* **96**, 424–434 (1983).
- Brady, S. T. *Nature* **317**, 73–75 (1985).
- Balczon, R. & Schatten, G. *Cell Motility* **3**, 213–226 (1983).

A molecular solution to the riddle of the giant panda's phylogeny

O'BRIEN *et al.*¹ predict "the existence of older Pliocene or even Miocene fossils of the ancestors of the giant panda". In fact, a putative ancestor was identified by Thenius² in 1979. This is *Agriarctos Kretzoi*, described on mandibular dental material from the late Miocene of Hungary. This ursid shows incipient molarization of the premolars apparently foreshadowing the condition in *Ailuropoda*.

This is independent support of the conclusions reached by O'Brien *et al.*, including their suggestion that *Ailuropoda* (with *Agriarctos*) may be assigned to a subfamily (*Ailuropodinae*) of the family Ursidae.

BJÖRN KURTÉN

Department of Geology,
University of Helsinki,
Snellmaninkatu 5, 00170,
Helsinki 17, Finland

1. O'Brien, J. O., Nash, W. G., Wildt, D. E., Bush, M. E. & Benveniste, R. E. *Nature* **317**, 140-155 (1985).
2. Thenius, E. *Anz. öst. Akad. Wiss.* **1979**, 67-68 (1979); *Z. Säugetierk.* **44**, 286-305 (1979).

Steens Mountain geomagnetic polarity transition is a single phenomenon

In their reinterpretation of the transition recorded in the Steens Mountain lavas¹, Valet *et al.*² suggest that the second phase (the 'rebound') occurred at a significantly later time than the first phase (the reverse-to-normal transition). In this way they attempt to separate what is one phenomenon into two separate phenomena, and assert that the second phase is a geomagnetic excursion unrelated to the preceding reversal. Several lines of evidence indicate that there was no significant hiatus in eruptive activity in the part of the lava section on Steens Mountain sampled by us and previously by Watkins³.

First, there are no soil horizons, weathered zones or interbasaltic sedimentary rocks or sediments anywhere in the sampled section, except some negligible baked earths (tuffs, perhaps) that we found in two restricted localities stratigraphically unrelated to the limits of either of the two transition phases. Moreover, the lavas are of rather uniform composition, and the only obvious lithological variability between successive lavas is the presence or absence of large platy labradorite phenocrysts up to 4 cm in diameter. The occurrence of these porphyritic lavas is unrelated to any aspect of the polarity transition.

Second, Gunn and Watkins^{4,5} made many complete chemical analyses of the

Table 1 Chemical groups⁵ and directional groups¹ for Steens lava analyses

Chemical group	Directional group
A, B	1-10
C	10-12
D	12-20
E	21-26
F	26-28
G	30-42
H	42-55

Table 2 Recalculated K-Ar ages

Watkins flow no. ^{1,5}	Directional group ¹	Age (Myr)
11	~4	15.5 ± 0.15
17	~10	15.4 ± 0.13
51	33	15.5 ± 0.28
61	43	15.4 ± 0.25
68	49	15.6 ± 0.21
70	49	15.3 ± 0.21

The uncertainty is 1 s.d. of the data.

Steens lavas. Table 1 shows how their chemical groups correspond to our directional groups. The first phase of the reversal begins with directional groups 43 and ends with group 30, and the second phase begins with group 28 and ends with group 15. The variations in chemical composition are uncorrelated with any aspect of the palaeomagnetic record, and several of the chemical boundaries occur within directional groups. The only possible peculiarity is that Gunn and Watkins analysed a high-alumina flow and a pyroxene-rich flow in chemical group F, which corresponds to the beginning of the second phase. These lavas, however, are perfectly correlated with the others, as demonstrated by an alumina variation diagram⁵.

Third, there is no discernible age difference between the top and bottom of the sampled section. We have recalculated

and averaged the K-Ar ages published by Baksi *et al.*⁶ and they are given in Table 2 (new decay constants).

Thus, the available field observations and geochemical and geochronological data directly contradict the suggestion of Valet *et al.*² that the two phases described in our article¹ may not belong to a single geomagnetic process.

C. S. GROMMÉ
E. A. MANKINEN*
M. PRÉVOT†
R. S. COE‡

* US Geological Survey,
345 Middlefield Rd,
Menlo Park, California 94025, USA

† Centre Géologique et Géophysique,
CNRS, and Université des

Sciences et Techniques,
34060 Montpellier Cedex, France

‡ Earth Sciences Board,
University of California,
Santa Cruz, California 95064, USA

1. Prévot, M., Mankinen, E. A., Grommé, C. S. & Coe, R. S. *Nature* **315**, 230-234 (1985).
2. Valet, J.-P., Laj, C. & Tucholka, P. *Nature* **315**, 217-218 (1985).
3. Watkins, N. D. *Geophys. J. R. astr. Soc.* **17**, 121-149 (1969).
4. Watkins, N. D. & Gunn, B. M. *Nature* **224**, 360-361 (1969).
5. Gunn, B. M. & Watkins, N. D. *Bull. geol. Soc. Am.* **81**, 1497-1516 (1970).
6. Baksi, A. J., York, D. & Watkins, N. D. *J. geophys. Res.* **72**, 6299-6308 (1967).

VALET, LAJ AND TUCHOLKA REPLY—In some of the reversal records that we have obtained from sedimentary sections in Greece (work partly published and presented at IAGA Praga 1985), we have also observed excursions occurring either before or after the transition. In all these cases the characteristics of the directional changes during both the polarity transition and the excursions are similar. The continuity of the sedimentary records also shows that the duration of the excursions never exceeds the time span of the transition.

In the Steens Mountain record¹, the first phase is quite similar to a sedimentary reversal record but, in striking contrast to our observations, the second phase is dissimilar. Furthermore, between the two phases the geomagnetic field has recovered high intensity and dipolar direction. The authors consider the two phases as a single phenomenon. In an attempt to reconcile sedimentary and volcanic records, we have suggested² that the widely different symmetry properties could indicate two unrelated different phenomena.

The existence of rebounds or oscillations associated with a reversal is an important problem which, as Prévot *et al.* suggest¹, may be crucial to our understanding of geomagnetic reversals. Clearly, what is needed is a criterion to establish when an excursion is connected to a reversal. The argument of continuity

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

suggested by Prévot *et al.* can reasonably be used as such a criterion. However, if the continuity can usually be well established in sedimentary records, this appears to us much more difficult in volcanic records. Furthermore, despite the additional information given here by Grommé *et al.*, we are not fully convinced that any of their arguments provide compelling evidence of a continuous mean extrusion rate on a timescale of less than 1,000 yr.

The rather uniform composition of lavas can hardly be used in this sense, as periods of constant chemical composition lasting up to about 150,000 yr have been observed, for example, in Etna³. We agree that the K-Ar dating results really indicate the same age for the top and bottom parts of the sequence. However, differences of about 0.3 Myr inside the same directional group (No. 49) are observed and if standard deviation is taken into account, the time span of age could reach 0.7 Myr.

Finally, the argument of the secular variation used by Prévot *et al.* to establish the linearity of timescale is also not convincing. Neither the number attached to the angular displacement of directions nor the stability of the secular variation over long periods of time is confirmed in longer records of secular variation, in lake sediments⁴ or by comparison of lake sediments and recent lavas⁵.

We are convinced that the data obtained by Prévot *et al.* provide a very impressive description of the geomagnetic field during a transition. What we are saying is that, in the present state of knowledge (or ignorance), part of the interpretation of this record might still be open to discussion.

JEAN-PIERRE VALET*
CARLO LAJ*

* Centre des Faibles Radioactivités,
Laboratoire mixte CNRS-CEA,
Domaine du CNRS,
Avenue de la terrasse,
91190 Gif-sur-Yvette, France

PIOTR TUCHOLKA†

† Laboratoire de Géophysique,
Université Paris-Sud,
Bât. 509, 91405 Orsay, France

1. Prévot, M., Mankinen, E. A., Grommé, C. S. & Coe, R. S. *Nature* **316**, 230-234.
2. Valet, J.-P., Laj, C. & Tucholka, P. *Nature* **316**, 217-218 (1985).
3. Tanguy, J. C. thesis, Univ. Paris (1980).
4. Creer, K. M. & Tucholka, P. *Phil. Trans. R. Soc.* **19**, 1106-1111 (1982).
5. Lund, S. P. *Geophys. Res. Lett.* **12**, 251-254 (1985).

Dust production in the Sahel

I WISH to make three points concerning the adequacy and consistency of the Sudanese visibility data used by Middleton in a recent analysis¹ relating the frequency of severe atmospheric dust occurrences to annual rainfall.

First, after 1974, the Sudan Meteorological Service started recording visibility

at 00.00 GMT in addition to 06.00, 12.00 and 18.00. Frequencies of poor visibility presented by Middleton for El Fasher in Fig. 1 (and for five Sudanese stations in Table 2) are therefore inflated for the years 1975-78 compared with previous years. Second, each observation of poor visibility does not necessarily refer to a separate dust storm event. Poor visibility at 06.00, 12.00 and 18.00 may be the continuation of a single storm whereas Middleton would categorize these as three separate storms. The resolution of the Sudanese data is insufficient to make this distinction. Again, the caption for El Fasher in Fig. 1 is misleading and the data are almost certainly inflated.

Finally, and most important, no discussion of the reliability of the visibility readings is included. Atmospheric visibility is a notoriously difficult climatic parameter to quantify. The Sudan Meteorological Service used the system of locating key landmarks at known distances from the observing station which are recorded either as seen or unseen by the observer at the designated times. I have observed ephemeral objects such as trees and electric lights being used at Sudanese stations. I would treat with great caution data collected in such a manner, where differences in procedure between successive observers can be crucial.

An inverse relationship between the recorded frequency of dust events and annual rainfall does appear to exist (subject to my first and second points) and the analysis is useful. I suggest, however, that climatic data from Sahelian countries be treated cautiously and their measurement and collection procedures carefully investigated and understood before they are analysed.

M. HULME

Department of Geography,
University of Salford,
Salford M5 4WT, UK

1. Middleton, N. J. *Nature* **316**, 431-434 (1985).

MIDDLETON REPLIES—I thank Dr Hulme for his interest. Dust storms are recorded by meteorological observers using fixed objects as visibility targets in all countries of the world where such phenomena occur, and such observations do present problems of observer error and involve an element of subjective decision making. As with all meteorological data there may also be changes in observational procedures. There are other problems inherent in interpreting data on the frequency of dust storms and dust storm days; I have highlighted some of the problems in a recent paper¹. Briefly, the statistics may conceal considerable variation in the areal and volumetric extent of dust storms and their duration. However, some of these problems are less serious when a

group of stations is used² to illustrate a particular trend in dust storm activity.

There may well be some inflation in the data for the Sudanese stations since 1974 as a result of a change in the frequency of visibility recording to which Hulme correctly refers, but the increasing trend that started before 1974 has continued to rise since that date and seems to be confirmed by studies in other countries of the Sudano-Sahelian belt. Note also that mean summer concentrations of dust at Barbados were again very high during 1982 and 1983 (ref. 3), although any comparison of dust storm frequency with dust mobilization or transport should be cautious because the scale of dust-raising events and the altitude to which material is lifted are of fundamental importance to the long-range transport of dust, and such information is not provided by dust storm frequency data for a particular station, as mentioned above.

N. J. MIDDLETON

School of Geography,
University of Oxford,
Mansfield Road,
Oxford OX1 3TB, UK

1. Middleton, N. J. *Search* **15**, 46-47 (1984).
2. Middleton, N. J. *Nature* **316**, 431-434 (1985).
3. Prospero, J. M. & Nees, R. T. *EOS* **65**, 837 (1984).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem

☐ Other (please specify) _____

☐ I am interested in sending my order by mail.

☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

UMI Article Clearinghouse

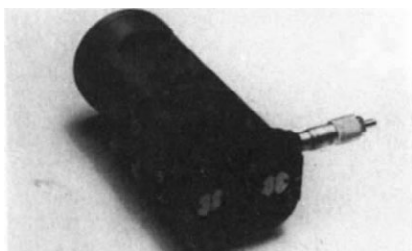
Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

New looks at spectroscopy

More spectroscopic equipment is becoming interfaceable with desktop computers — and even the Bunsen burner is going high-tech.

• Weighing only 2 pounds and compact in size, the Spectra-Physics IR minimonochromator system (Model SP5155) covers the 2.7–30 μm region. The optical configuration is that of an Ebert monochromator with grating blaze options. Standard interface optics packages are available for interfacing the unit with Spectra-Physics tunable diode lasers and the SP5800 laser source assembly. The minimonochromator is a product of the Laser Analytics Division of Spectra-Physics.

Reader Service No. 100.



IR minimonochromator system SP5155.

• The new RC, high efficiency recirculating coolers from New Brunswick provide a continuous and reliable capacity from 750 to 58,500 watts and allow fluid temperature control between -40°C to $+75^{\circ}\text{C}$ with accuracy of 0.5°C . Coupled with one of the variety of built-in pumps available these recirculating coolers are ideal for the heat removal requirements of most laser systems. The optional digital controller and IEEE/488 bus allow computer control of the system for complete automation.

Reader Service No. 101.

• The Varian VXR series NMR spectrometers are superconducting Fourier transform instruments functioning at 100, 200, 300, 400 and 500 MHz. All critical spectrometer functions are under computer control, and components are modular for high reliability, ease of maintenance and for easy upgrading. For data manipulation and experiment management, the multiple processor-based VXR 4000 data station includes a 68000-based computer and a Winchester capacity of up to 160 Megabytes. The VXR 4000 may use optional networking capabilities including Ethernet, Varian's magNET and Serial Communications software. For data processing and display, the VXR 4000 may be purchased as an independent work station.

Reader Service No. 102.

These notes are based on information provided by the manufacturers. To obtain further details about these products use the reader service card bound inside the journal.

• A self-contained extract system for atomic absorption spectrophotometers is now available in kit form from Mason Nordia. The extract system will protect the operator against toxic fumes and protects the instrument against corrosive vapours. To reduce the risk of heat gain in the laboratory, the kit has specific capability to overcome problems of heat concentration where the flame generates 3kW or more of heat.

Reader Service No. 103.

• The latest catalogue on optical table and isolation systems from Photon Control of Cambridge includes: large high performance metal honeycomb and synthetic granite optical tables designed to meet the needs of the laser and optical scientist; up-to-date modular designed vibration isolation systems; and special purpose tables and vibration isolation systems on which to mount precision instruments and sensitive equipment. Use the readers service to obtain a free catalogue.

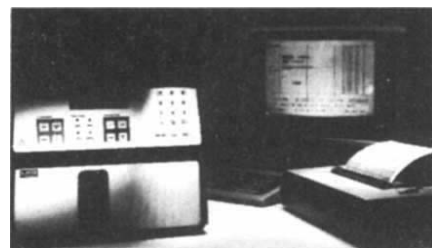
Reader Service No. 104.

• Blak-Ray UV intensity meters are designed for monitoring the efficiency and uniformity of UV light sources. The J-225 model measures in the short-wave range from 220 to 280 nm with peak sensitivity at 254 nm. Long-wave readings are made with the J-221 model with a range from 300 to 400 nm and a peak sensitivity at 365 nm. Readings are given in $\mu\text{W cm}^{-2}$. Sensors plug directly into the top of the meter housing. Readings in hard-to-reach locations are made using a 4-foot cable to connect the sensor to the meter. Each model features a A and B Scale for low or high intensity readings. A selector switch chooses the scale, and a snap-on accessory filter provides true UV readings with infrared present.

Reader Service No. 105.

• A new infrared detector is available for the Rofin-Sinar optical spectrum analysis system. The detector enables the system to rapidly scan optical spectra in the 1.5 to 4.5 μm wavelength region with a resolution of 20nm. The resulting spectral data (10 scans per second) may be displayed in real time on a standard laboratory oscilloscope or it may be processed by an Apple or BBC computer before being displayed on a monitor or printer. The detector unit, model RSO 6115, uses a PbSe element as the optical transducer and it is supplied with a stabilizer bias power supply, a signal pre-amplifier, and all the necessary mounting hardware for easy connection to the RSO 6140 scanning monochromator.

Reader Service No. 106.



LKB's PC-interfaceable liquid chromatography.

• LKB-Produkter AB of Sweden has developed extended PC software for its diode array detector for use in liquid chromatography. This new Wavescan software makes it possible to automate diode array detection even further, and offers a significant increase in post-run analytical productivity. Designed for the PC, XT, AT or any IBM compatible, Wavescan programs use a simple menu technique. With instant access to isograms, chromatograms and spectra, it is possible to obtain considerably more detailed information on sample identity and purity.

Reader Service No. 107.

High-tech 'Bunsen'

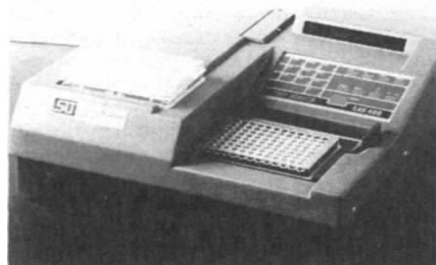
THE conventional Bunsen burner may be a thing of the past. From Technomara of France comes the Fireboy, with an electronically controlled gas flame. Exotic features include a provision for a control pedal operated by the foot to free both hands for experimental procedures, and accurate control of the flame temperature. *Reader Service No. 108.*



• Questek excimer lasers, available from Lambda Photometrics, provide high power UV pulses (40–750MJ) at any one of a number of fixed wavelengths. A new 20-page brochure is now available, giving full technical details, applications data and suggestions on selecting the right model. *Reader Service No. 109.*

• Australian Monoclonal Development (AMD) now offers a T-cell panel for enumeration of total T cells and T-cell subsets using the indirect immunofluorescence method. The AMD T-cell panel includes three monoclonal antibodies: T4, T8 and T11 equivalents, plus FITC anti-mouse Ig conjugate. Future releases will include a conjugated panel of the same antibodies. These reagents are currently being used routinely to determine T4/T8 ratios in suspected AIDS patients and also to observe T cell imbalances in other immunological diseases. AMD will soon be marketing two other tests for clinical use. The Monohaem kit is a reliable immunoassay for occult blood in faeces and is based on a monoclonal antibody which is specific for human haemoglobin. The AMD Gono-kit uses a panel of monoclonal antibodies which covers all strains of *N. gonorrhoeae* for definitive identification by direct immunofluorescence on clinical specimens. *Reader Service No. 110.*

• New from SLT Labinstruments, the SLT Easy Reader plate photometer is designed with enzyme-linked immunosorbent assay in mind. The Easy-Reader is ideal for low volumes, combining fast, accurate measurements with sophisticated software for data evaluation. The Easy



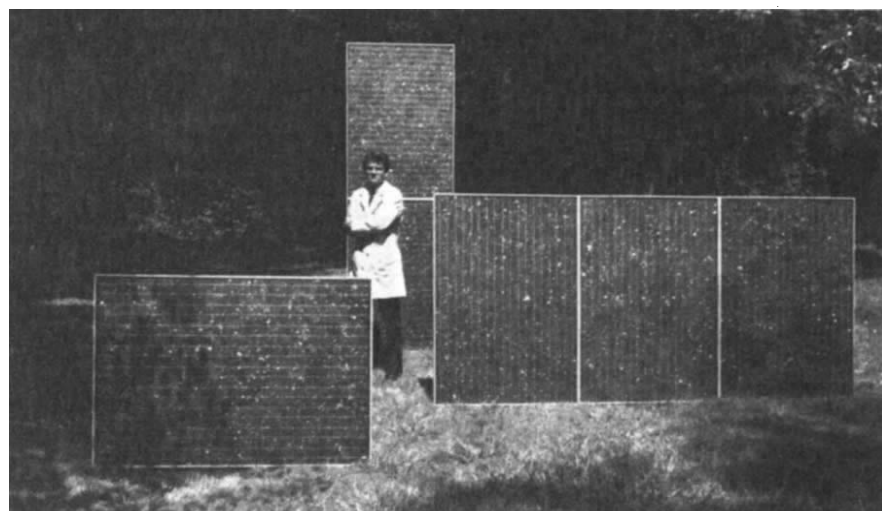
SLT's Easy Reader.

Reader Measures the whole plate with one filter in 20 seconds and for dual wavelength measurements it takes only 35 seconds.

Reader Service No. 111.

• Hewlett-Packard's HP 9000 computer with enhanced mass-spectrometer software will enhance the performance of the HP 5970B mass-selective detector (MSD), HP 5995C GC/MS, and HP 5988A GC/MS and LC/MS systems. The workstation combines enhanced GC/MS, LC/MS and MSD software with the new HP 9000 Series 300 technical desktop computer, which is designed specifically for instrument control and scientific-data handling.

Reader Service No. 112.



The largest solar panels on the market, from Solarex. The modules are available as 145, 290 and 435 watt units and achieve a high efficiency of 12-4%. *Reader Service No. 113.*

• Now available in the UK from Baird & Tatlock, the Cecil range of UV/visible spectrophotometers provides three single-beam models, each with a choice of four systems. The three models available have a wavelength range of 190 to 990nm and a wavelength accuracy within 1nm. The models are: CE2202, measuring transmittance and absorbance on a meter scale, ideal for teaching or routine work; CE2272, with a linear scale for transmittance, concentration and wide scale expansion ranges for absorbance; and CE2293, with a digital scale for absorbance transmittance and concentration. It also has the option of computer interfacing, data handling and an autozero facility. *Reader Service No. 114.*

• Rigaku manufactures two similar X-ray fluorescent spectrometers — the 3070 and the 3071. The 3070 spectrometer has normal optics; the 3071 features inverse optics. The 3071 allows for greater ease of liquid sample handling and analysis, whilst retaining the same features as the 3070 model. Both are menu-driven for ease of operation, with an internal-external sample chamber for continuous feed analysis, spectrometer status display, full data reduction capabilities, spectrum marker elemental identification and computer controlled pulse height analyser. *Reader Service No. 115.*

• New options on the Perkin-Elmer 1700 Series of Fourier transform infrared spectrometers include an extended near IR range. The standard model 1750 now scans to 5,000 cm^{-1} in the near infrared, a region of the spectrum particularly useful for quantitative analysis in the polymer industry. Model 1710 scans to 4,400 cm^{-1} , and this wavelength range is extended to 5,000 cm^{-1} when the spectrometer is used with the model 3600 data station. And a caesium iodide beamsplitter is now available as an alternative to the standard KBr beamsplitter. Fitted with this, the model 1710 scans from 4,300 cm^{-1} to 220 cm^{-1} and model 1750 from 5,000 cm^{-1} to 220 cm^{-1} .

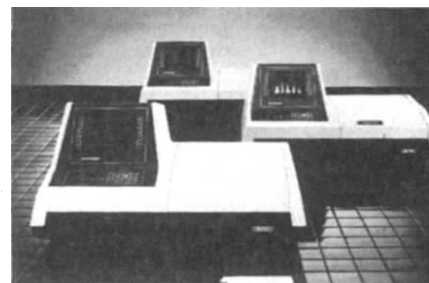
Reader Service No. 116.

• VG Isotopes, specialists in isotopic and elemental mass spectrometry, has announced a new analytical research service for scientists, researchers and technologists. VG Isotopes Analytical Service Group has at its disposal an array of techniques including inductively coupled plasma mass spectrometry (ICP-MS), glow discharge mass spectrometry (GDMS) and thermal ionization mass spectrometry (TIMS). The analytical service should be particularly attractive to those who would not otherwise have access to such advanced instrumentation, or need only occasional access to such equipment.

Reader Service No. 117.

• The Gilford Response range of UV/Vis spectrophotometers from Ciba Corning has now been extended to three separate models. Common to all models is: an autoranging photometer; transverse sample compartment; high resolution graphics; data storage facilities and a 16-bit resident computer. The basic Response is supplied complete with a single cell holder and programs for: wavelength scanning; kinetics; time scanning; multi-wavelength analysis; standard curves and sample read. The second unit has an automatic six-position cell holder as standard and has the capability of adding gel scanning, rapid sampling and a six-position thermostet cuvette holder for electronic control of sample temperature. The most advanced Response has two disk drives for archival data storage, together with extended software applications.

Reader Service No. 118.



Response spectrophotometers.

PRODUCT REVIEW

ADVERTISEMENT

• The new Gallenkamp Visi-Spec is a low-cost spectrophotometer for measurements in the visible spectrum. It is suitable for general purpose measurements where cost and reliability are important. The complete optical system is microprocessor controlled. There are only four push button keys and it automatically goes through a three-point wavelength calibration check every time it is switched on. Standard 10-mm cuvettes or test-tubes are accepted.

Reader Service No. 119.



Gallenkamp's Visi-Spec.

• The Oriel Scientific DRTA-9 integrating sphere accessory for the Perkin-Elmer Lambda 9 UV-Vis-NIR spectrophotometer is a six-inch diameter sphere which can collect either scattered light from turbid solutions or reflected light from solid samples. This accessory expands the usefulness of the Lambda 9 beyond its standard application for homogeneous, translucent solutions and makes it possible to perform quantitative analysis of solid materials used in solar and optical component research. The analytical chemist and life scientists are offered qualitative analysis of turbid solutions, NIRA and Kubelka-Munk functions without extensive sample preparation. The DRTA-9 fits entirely inside the Lambda-9 sample compartment cavity and is easily installed by the system operator.

Reader Service No. 120.

• Kontron Instruments' Uvikon 860 UV/Vis spectrophotometer offers a combination of intelligent electronics and proven optics, with the intelligent electronics making full use of a video display and RAM curve memory for entering parameters and post measurement calculation. Each parameter is presented to the operator, together with its possible value. The Uvisoft keys which are software defined function keys appear on the video display



The Uvikon 735LC.

allowing the Uvikon 860 to be very user friendly. In addition to the normal measurement modes, 'fix' and 'scan' the Uvikon 860 has programs for enzyme kinetics, end-point analysis and absorbance versus time. By using a personal computer the operator also has access to the internal commands of the Uvikon 860 and personalized programs can easily be written to suit the operator's own application. In addition, the Uvikon 860 allows complete sets of method parameters to be stored in permanent memory, ready for fast set up of routine conditions.

Reader Service No. 121.

• Cuvette sample packs are being offered by Elkay Products Inc. free of charge to laboratories wishing to evaluate the company's six different disposable cuvettes. Precision moulded in UV-inhibitorless polystyrene or acrylic copolymer, the cuvettes are capable of averaging 70% light transmittance at 340 nm. Each cuvette is packed in particle free, shrink wrapped reusable polyfoam trays.

Reader Service No. 122.



Free cuvettes from Elkay.

• Beckman's data capture software for use with DU-7 and DU-50 UV/Vis spectrophotometers is compatible with the IBM PC series of desk-top computers and features colour graphics with easy colour selection. The software allows single-wavelength, scanning, kinetics and gel scan data to be transferred to the computer and conveniently stored in labelled files on standard disks. With the DU-7 spectrophotometer stored data can be transferred bidirectionally so that data from disk may be viewed and manipulated on the graphical video display of the DU-7. With the DU-50 series stored data can be displayed quickly on the computer's monitor with overlay of curves in different colours, and scans can be added and subtracted to give accumulated results or difference spectra. All data are stored in a standard file format which is directly compatible with such popular general software as Lotus 1-2-3. With Lotus, data can be entered into a spreadsheet presentation for carrying out complex calculations and data manipulation.

Reader Service No. 123.

**Geological Sciences Series,
Vol. XIX Stanford University**

New Publication.

Basic-Hypo; A Basic Language Hypocenter Location Program User's Guide

**Jorge Mendoza
Dale Morgan**

An efficient BASIC language computer program for the location of hypocenters of local and regional earthquakes, is implemented on a 5 1/4" diskette.

The publication includes the source code and it can be typed in computers, like Commodore 64 (TM), Radio Shack TRS 80 (TM), Hewlett-Packard Series 80 (TM), Apple II family (TM), IBM PC (TM) and all PC compatibles and any other computers that run Microsoft Basic (TM).

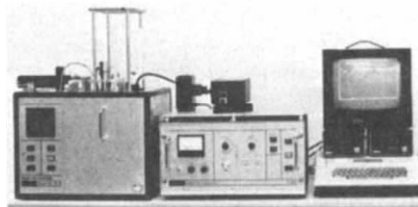
The diskette included in the publication is MS DOS compatible.

Package includes a 32 page report, soft covered with a ready to use 5 1/4" diskette. Priced at \$25 each. Please include 7% tax for California requests.

—ORDER FROM—

**Stanford University
Department of Geophysics, SRP
Stanford, California 94305
Att. Margaret Muir**

• For those involved in the study of rapid reaction intermediates, Hi-Tech has introduced the SFL-43 stopped-flow/multimixing flow module as an additional capability to the SF-4 spectrophotometer system. This flow module can be used in two modes in the study of reaction intermediates or just simply for high performance stopped-flow work. In the sequential mixing mode, two stable solutions are mixed to form a short-lived intermediate which in turn is mixed with a third reagent after a pre-set delay as short as 5 ms. The second reaction is studied by stopped-flow spectrophotometry. The preparative quench flow made enables a fast chemical or biochemical reaction, that



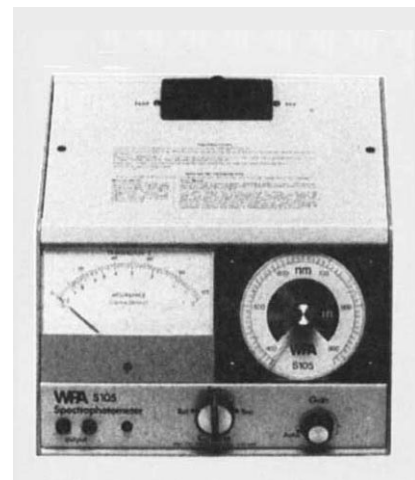
SFL-43 augments the SF-4.

lacks a suitable chromophore, to be analysed by an intrinsically slow technique such as NMR spectroscopy. Initiated by mixing, the reaction is stopped after a pre-set delay by mixing with a quench solution and the product is then collected for analysis. By varying the delay a series of values may be plotted to determine the reaction rate. The flow circuit is automatically purged by solvent before the run, thus eliminating cross-contamination.

Reader Service No. 125.

• Bio-Rad's new Model 1745 fixed-wavelength UV monitor is a highly sensitive detector suitable for both HPLC and open column applications. In standard form, the instrument is able to monitor UV absorption at either 280 nm or 254 nm — the most popular wavelengths for protein, pharmaceutical and nucleic acid applications. For peptide studies a low-cost zinc lamp can be added to provide detection at 214 nm.

Reader Service No. 126.

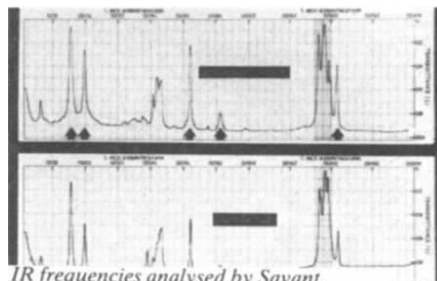


WPA S105 spectrophotometer.

• The S105 and S107 spectrophotometers from Walden Precision Apparatus are designed to cover the UV and IR ranges, operating between 380 and 920 nm. A slow motion wavelength selector control with a large scale and pointer is incorporated in these instruments. The scale is marked at 5-nm intervals and the approximate colours of the visible spectrum are shown to ensure accurate wavelength selection. The symmetrical optical system achieves a bandwidth of less than 5-nm and the illumination is provided by a special tungsten filament lamp which gives a continuum of emission from about 300 nm to well into the infrared spectrum. Both spectrophotometers accept standard glass or plastic 10 mm cuvettes which are shielded from extraneous light. A quick

action rotary control allows the change from reference to sample to be made easily. The S107 is fitted with a fan cooling unit which allows the cuvettes to be maintained at closer temperature tolerances. In addition to manual operation, the S105 can be used in a semi-automatic mode. It can also be provided with a scanning facility by the addition of a small electric drive unit and, when coupled with a chart recorder, this instrument can make fully automatic scans.

Reader Service No. 127.



• An audiovisual training program on the interpretation of infrared spectra, written by J.H. van der Maas and E.T.G. Lutz has been published by Savant Audiovisuals and John Wiley. Composed of eight lessons which systematically cover the major IR group frequencies, the package contains 123 slides, four tapes, 10 course books, an answer book and 100 correlation tables. It is illustrated with high quality spectra of known compounds and unknown materials for student practice. This educational package (IR 1000) complements Savant's other infrared audiovisual programs which are concerned with techniques, instrumentation and applications. They cover infrared quantitative analysis (IR-101), solid sample handling for infrared spectroscopy (IR-102), infrared internal reflectance spectroscopy (IR-103) and computerized infrared spectroscopy (IR-104).

Reader Service No. 128.

• Burleigh Instruments, Inc. has introduced two new mirror sets which extend the spectral range of the company's SA-800 and SA-200 Series optical spectrum analysers. The new mirror sets cover the important wavelength regions of 800–900 nm and 1.06 μm. Burleigh spectrum analyser systems are modular and feature an x, y, theta and phi adjustable mount for complete control over the alignment process. Each system is delivered with spectrum analyser, controller, detector, beamsplitter, mount and bases. The analysers are being used extensively as a diagnostic for mode-locked YAG systems.

Reader Service No. 129.



Burleigh's RC-46 controller with accessories.

ADVERTISEMENTS

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID

Our new commercial LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized

- UNILAMELLAR OR
- OLIGOLAMELLAR OR EVEN
- MULTILAMELLAR LIPOSOMES

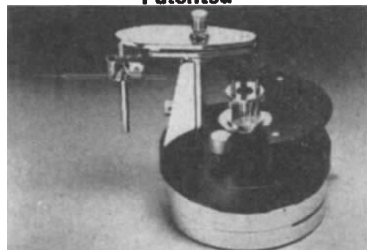
depending on the used experimental conditions. Liposome size is selectable between 25 and ca. 600 nm diameter up to 1000 nm for multilamellar liposomes. Sample volume is normally 5–10 ml or up to 200 ml if used in combination with an additional instrument. High encapsulation efficiency for hydrophilic and/or lipophilic agents. Lipid concentration can be up to 300 mg/ml. Lipid loss during preparation is extremely low: down to 0.1%.

BIANORM - Geräte

POB 126, D-8000 Munich 65, FRG

Reader Service No.36

SINGER MICROMANIPULATOR MK III Patented



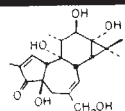
Perfect control in 3 Dimensions from a single joystick. Continuous variable reductions to suit all magnifications. No after sales maintenance needed.

Details from:

Singer Instrument Co. Ltd.,
Treborough Lodge Roadwater,
Watchet Somerset TA23 0QL
England. Tel. (0984) 40226 Telex:
337492 Comcab G.

Reader Service No.48

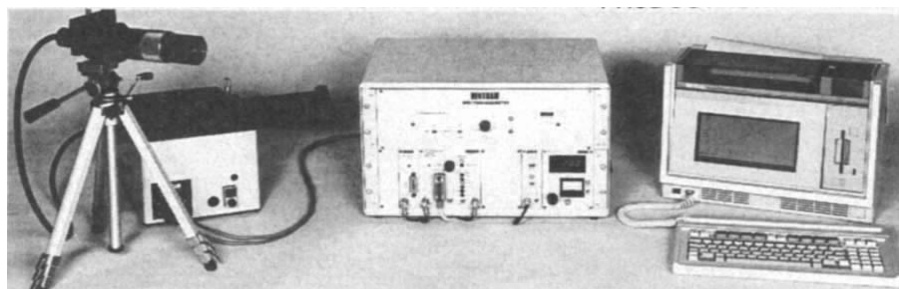
PHORBOL ESTERS



LC SERVICES CORPORATION

165 New Boston Street
Woburn, MA 01801
(617) 938-1700

Reader Service No.11



The Bentham spectral analysis system with software.

• Bentham's new modular spectral analysis system can be used over the wavelength range from 200nm to 30 μ m. The system can as easily be configured for accurate colour measurement of displays as it can be for measurement of transmission and emission in the ultraviolet and infrared regions. Light sources, detectors and electronics including lock-in amplifiers are available, while system control is facilitated by the use of the IEEE/488 interface bus (GP-IB, HP-IB). Popular applications include colour measurement, for which calibrated standards are supplied, measurement of spectral loss in fibre optics and monitoring of plasma etching processes. A software package available for HP-85 and some other popular desk-tops perform general spectral measurement and fibre loss.

Reader Service No. 130.

• Polyspec, is a new microspectrophotometer from Reichert-Jung. It is a modular system comprising a basic system for measuring transmittance (extinction), reflectance or fluorescence which can be upgraded to microspectral analysis or to an integrated complete microscope measurement system with automatic scanning control. Providing high sensitivity and reproducibility, this new system permits observation of the measuring diaphragm during measurement by employing the synchronized dual chopper principle. The Polyspec has a choice of two operating modes — either the HP (high precision) chopped mode taking 140 measurements per second, or the HS (high speed) d.c.-mode at 1,000 measurements per second. The control unit employs 2 Z-80A microprocessors, it can be fitted with up to eight plug-in dedicated applications program modules, and has a liquid crystal display capable of graphics presentation. The Microspectrophotometer may be interfaced to a computer for programming and data processing and to a printer-plotter.

Reader Service No. 131.



Reichert-Jung's Polyspec.

• Philips Analytical has developed two new low-cost ultraviolet/visible spectrophotometer data systems based on the PU8800 series of instruments. The PU8800 DS system, designed for scanning work, consists of the high performance double beam PU8800 spectrophotometer, a computer customized for the laboratory environment, and a scanning software pack. The software permits spectral display, overlay and expansion using a zoom facility. Spectra, each containing up to 7,000 data points, can be stored on disk, together with optimized display parameters and PU8800 scan conditions. The second system, PU8800 DSK, is specially designed for kinetics and comprises a PU8800 spectrophotometer, computer, six-cell changer and an advanced kinetics software pack. Features include absorbance against time plots and rate calculation with linearity check. Printouts of rate, activity and purity can also be obtained. K_m and V_{max} values can be calculated by either a direct linear plot or regression analysis, and the plots and results displayed in a few seconds.

Reader Service No. 132.



The PU8800 UV/Vis spectrophotometer.

• A completely new analytical transmission electron microscope from Carl Zeiss (Oberkochen) incorporates a fully integrated electron loss spectrometer. This gives the Zeiss EM902 the capability of performing Electron Spectroscopic Imaging (ESI), providing significant improvements in contrast, and high resolution elemental mapping over conventional techniques. Using ESI it is possible to analyse and image light elements not accessible to X-ray microanalysis with sensitivities and resolutions not even theoretically attainable by X-ray spectrometry.

Reader Service No. 133.

• The latest issue of Finnigan MAT publication Spectra covers clinical applications of mass spectrometry and includes articles on organic acidurias in children, use of GC/MS in clinical toxicology, eicosanoid mass spectrometry and steroid profiling by GC/MS.

Reader Service No. 134.

• The Kratos Analytical MS80 is a high performance medium resolution mass spectrometer designed for both routine operation and advanced research problems. It features double focusing forward geometry that incorporates hexapoles, with advanced source construction and a novel magnet design that gives outstanding sensitivity maintained across the range of resolving power. Exceptional capabilities are available in quantitative selective ion monitoring, at picogram levels and below, of materials in complex mixtures. A comprehensive range of inlet systems includes gas chromatography, with facilities for packed and capillary column work. Fully programmed digital logic and display units give push-button control of major instrument functions. The fast scanning RF magnet is of particular use in capillary column GC/MS, both in terms of its fast scan speed and also in its speed of switching in magnetic selectors ion monitoring mode. The MS80 can be provided with an LC/MS inlet utilizing the Thermo-spray mode of operation.

Reader Service No. 135.

• A new mass spectrometer for automatic ^{15}N analysis has been introduced by VG Isogas, the ANA-SIRA (for Automatic Nitrogen Analyser-Stable Isotope Ratio Analyser) mass spectrometer system. The ANA-SIRA can automatically analyse both solid and liquid samples for ^{15}N and total N with sample sizes as low as 5 μg N at natural abundance. This combination of sample preparation and analysis enables fast (5 min per sample), reproducible (± 0.0005 atom % ^{15}N), and automatic analysis in the fields of biomedical, agricultural and environmental research. An ANA Interface upgrade kit is available for all VG Isogas mass spectrometers.

Reader Service No. 136.



LeCroy 9400 Digital Oscilloscope

• Fast logging of waveforms data is particularly useful when signal availability is short, for instance in electron spin resonance experiments and laser spectroscopy measurements. A segmentable 32K bit memory in each channel of the new LeCroy 9400 Digital Oscilloscope enables logging of successive events on trigger command. This allows very fast acquisition of waveforms, since no transfer of digitized data to external devices such as a computer or disk is needed. Up to 250 waveforms of 125 words can be logged well within 50 ms at high time-base speeds in the acquisition memories, or 125 waveforms of 250 words within 25 ms, or as fast as 8 waveforms of 2,500 words within 2 ms.

Reader Service No. 137.

Graduate mobility in the United Kingdom

from Richard Pearson

Increasing interest is being paid in the United Kingdom to longer-term graduate careers; several cohort studies following graduates from 1980 and later years are now under way or are being planned.

Job prospects and employment patterns of new qualifying graduates are seen by many as an important indicator of higher education's "success" and relevance to industry. Others argue that they say nothing about lifetime returns from higher education which should not be exclusively job-related. Nevertheless, by providing an easily quantifiable indicator, trends and comparisons based on the first destination statistics will be made. While they provide a valuable barometer of initial labour market success, they are not a definitive statement of longer-term employment prospects. A typical working life lasts forty years or more and in the United Kingdom there are more than six million job changes each year. Evidence about employment status six months after graduation can provide only a starting point for assessing graduate careers.

In providing a broader context, a recently published study¹ of graduate mobility based on the experiences of 280 em-

ployed graduates five years later. This may be in part explained by the low recruitment levels in this sector in the early 1980s resulting in a shortage of experienced graduates three to five years later which in turn encouraged mobility between employers. By contrast, only 26 per cent of those recruited by chemical companies in 1980 had left their initial employer in the following five years. More general evidence suggests that the retention curve flattens after five years or so, and that after 10 or 15 years probably 30 per cent of graduates are still with the same employer.

These studies show only general trends. The best source of data on the mobility of individual graduates remains the now dated study by the Department of Employment of 12,000 students and their careers and mobility over the seven years to 1977. That study showed that men graduates were more mobile than women graduates, both between their occupations and type of work, but that women were more likely to move between employment sectors. The main motivation of graduates in changing jobs was the "job interest" and "level of responsibility offered" but the prospects of a good salary and promotion assumed increased importance as their careers developed. Women's salaries were about 20 per cent below those of men after seven years, a difference only partly explained by the higher incidence of part-time work among women. The highest salaries were, in 1977, paid to graduates in social studies or engineering, the lowest to arts graduates. Managerial jobs in industry and commerce tended to be the best paid, while school teachers were among the lowest paid. It was also found to be less rewarding financially seven years after graduation to have studied and gained a postgraduate qualification than to have entered employment directly and gained work experience. This study is due to be repeated next year, based on the cohort graduating in 1980.

The most recent major survey of graduates' early careers has been that being undertaken by the Council for National Academic Awards (CNAA) which is surveying 10 per cent (2,600) of graduates who received CNAA degrees in 1982. The students were surveyed in 1983, 1984 and 1985, allowing a three-year profile to be drawn up of their early careers. The results of the 1983 survey reveal a complex pattern of transition over the first year into the labour market. They show significant numbers entering, for

example, part-time or temporary employment and one in six changing their job within the first year. In many of the "vocational" subjects such as pharmacy, nursing and business studies, the graduates were in their preferred job, but 40 per cent of all graduates and 60 per cent of scientists were looking for "something better". Nearly half the scientists felt over-qualified for the job they were then doing and one in six scientists had already had two or more jobs in their first year after graduation. A further 41 per cent had considered changing jobs. Nearly a year after graduation, the average salary for a scientist in 1983 was £4,867, considerably below that of engineers (£6,358) and maths/computer scientists (£6,354) but similar to

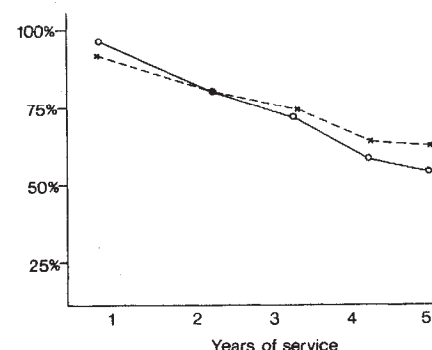


Fig. 1 Graduate retention. Taken from ref. 2.

that for social scientists (£4,865) and just ahead of the arts and humanities graduates (£4,672). The average salary across all subjects was £5,058. More than one in four of the CNAA scientists was unemployed ten months after graduation which is rather more than at the six months stage.

For graduates, like all other first time entrants into the labour market, the first job is only the first part of a career, the early years of which are likely to see several changes of job and possibly of employer. The first year for many can be a particularly difficult year of transition, especially, as in 1982, in times of a depressed labour market. The first destination statistics provide one snapshot, and give a limited performance measure; they do not however provide a complete picture of life after higher education.

1. Parsons, D. *Graduate Recruitment and Retention 1980-84*. HMS, 1985.
2. CNAA Graduate Survey. Interim Report, 1985.

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.

Graduate retention 1974-79 and 1980-84

Retention dates	Year of cohorts*				
	5th	4th	3rd	2nd	1st
1974-79 cohort	54	59	71	84	96
1980-84 cohort	58	66	72	84	94

* 1st = most recent that is 1979 and 1984 respectively.

ploying organizations showed that nearly 40 per cent of graduates recruited in 1980 were no longer with the same employer five years later. As Fig. 1 shows, there is a fairly standard decay curve, with 34 per cent of the 1981 cohort and 16 per cent of the 1983 cohort having left their initial employer by 1985. In addition, most graduates who have stayed with a single employer will also have changed jobs over this period and some will also have moved geographically. What is particularly interesting about the 1985 study is the comparison that can be made with the results of a similar study in 1980. The time span for each was five years, and each covered a similar economic cycle from a low to a high point of economic activity. Despite the fact that the recession of the 1980s and the contraction in job opportunities was far more severe, retention rates followed remarkably similar patterns (Fig. 1). Despite this consistency over time, there were differences between employers from different sectors. The sector with the highest loss was engineering, where 60 per cent of the 1980 graduates were no longer with their